



دانشگاه صنعتی شاهرود

دانشکده علوم ریاضی

گروه ریاضی کاربردی

پایان نامه کارشناسی ارشد

عنوان

آزمون گروهی

نگارش

سمیه خواجه

اساتید راهنما

دکتر نادر جعفری راد

دکتر میثم علیشاهی

شهریور ۱۳۹۱

آزمون گروهی

سمیه خواجه

۲۲ مهر ۱۳۹۱

کلیه حقوق اعم از چاپ و تکثیر، نسخه برداری، ترجمه، اقتباس و ... از این پایان نامه برای
دانشگاه شاهرود محفوظ است.
نقل مطالب با ذکر مأخذ آزاد است.

صفحة تصویب نامه توسط هیأت داوران (فرم پیوست ۳ یا ۴ با امضای اصل هیأت داوران مورد قبول است)

قدردانی

شکر خدا را که هرچه طلب کردم بر منتهای حکمت او کامران شدم. در ابتدا وظیفه خود می‌دانم که از زحمات خانواده عزیزم بخصوص پدر و مادر مهربان و دلسوزم که آرامش روحی و آسایش فکری فراهم نمودند تا با حمایت‌های همه جانبه‌شان این مجموعه را به پایان برسانم کمال سپاسگزاری را ادا نمایم.

همچنین لازم است از زحمات استاد راهنمای بزرگوارم جناب آقای دکتر میثم علیشاهی که با همت بی‌شائبه‌شان گلشن سرای علم را در درونم بارور ساختند تقدیر و تشکر نمایم.

در خاتمه لازم است مراتب تقدیر و تشکر را از آقایان دکتر نادر جعفری‌راد، دکتر علی طاهرخانی و دکتر جعفر فتحعلی اعلام نمایم.

چکیده

در این پایان نامه، آزمون گروهی ایستا را ارائه می دهیم و همچنین سه نوع از ماتریس های دودویی را به عنوان ابزار مهم در فهم و ساختن آزمون گروهی ایستا تعریف می کنیم. نشان می دهیم که آزمون گروهی ایستا برای شناسایی حداکثر دو بیمار از میان 2^t نفر نیاز به حداقل $2t$ آزمون دارد. از گراف ها در ساختار آزمون گروهی استفاده می کنیم و همچنین نشان می دهیم که تحت شرایطی روی اندازه گروه می توان با استفاده از گراف پترسن تعمیم یافته آزمون گروهی ایستای بهینه ای را ساخت. در ادامه الگوریتم های آزمون گروهی ایستا را ارائه می دهیم و آنها را به حالت خطا آزاد، وجود پوچگرها، مدل هم تافت و مدل آستانه تعمیم می دهیم و همچنین ساختارهای ترکیبیاتی جدیدی به همراه کاربردهای آن بر روی آزمون گروهی ایستای بهینه با وجود پوچگر را ارائه می دهیم. در انتها به این پرسش پاسخ می دهیم، که چه موقع در یک آزمون گروهی ایستا، آزمون انفرادی بهینه است؟ فرض می کنیم که برای ثابت d ، $N(d)$ بزرگترین مقدار n باشد که برای آن آزمون انفرادی بهینه است. نشان می دهیم که برای $d = 1, 2, 3, 4$ ، $N(d) = (d + 1)^2$ است.

واژه های کلیدی: آزمون گروهی غیر انطباقی، ماتریس مجزا

پیشگفتار

فرض می‌کنیم در بین n نفر حداکثر d تا از آن‌ها بیمار هستند. می‌خواهیم با انجام آزمایش افراد بیمار را شناسایی کنیم. از طرفی اکثر آزمایش‌ها هزینه‌بر هستند و همچنین با توجه به فسادپذیری خون تعداد زیاد آزمایش برای شناسایی افراد بیمار امکان‌پذیر نیست. اگر S یک زیرمجموعه از کل افراد باشد می‌توانیم همه‌ی افراد واقع در S را یکباره مورد آزمایش قرار دهیم و اگر نتیجه آزمایش منفی باشد نتیجه بگیریم همه‌ی افراد S سالم هستند. لذا هدفی که در این راستا مطرح می‌شود یافتن یک الگوریتمی است که با کمترین تعداد آزمایش گروهی بتوانیم افراد سالم و بیمار را شناسایی کنیم.

یکی از این الگوریتم‌ها آزمون گروهی ایستا است. در این الگوریتم ابتدا خون افراد را به چند گروه تقسیم می‌کنیم و سپس بطور همزمان روی تمام گروه‌ها آزمایش انجام می‌دهیم بعد از انجام آزمایش ما باید قادر باشیم تمام افراد بیمار را شناسایی کنیم. از آنجا که در الگوریتم ایستا همه‌ی گروه‌ها باید در ابتدا مشخص باشند به‌وضوح در این روش مسئله‌ی اصلی چگونگی گروه‌بندی کردن نمونه خون افراد است. به این الگوریتم می‌توان یک $(1, 0)$ -ماتریس نسبت داد به‌طوری‌که هر سطر آن متناظر با یک گروه و هر ستون آن متناظر با یک شخص باشد. حال باید ببینیم چه ماتریس‌هایی و با چه خاصیت‌هایی یک الگوریتم ایستا را نتیجه می‌دهد.

در فصل اول این پایان نامه به تاریخچه آزمون گروهی، تعاریف و مفاهیم مورد نیاز بصورت اجمالی اشاره می‌شود. در فصل دوم این پایان نامه سه نوع از $(1, 0)$ -ماتریس‌ها که متناظر با الگوریتم ایستا هستند، مورد مطالعه قرار گرفته‌اند و همچنین قضیه‌هایی در چگونگی ارتباط این ماتریس‌ها با یکدیگر بیان شده است.

در ادامه این ماتریس‌ها را به حالت خطا آزاد تعمیم می‌دهیم و مشاهده می‌شود که با فرض وجود خطا در نتایج آزمایش گروه‌ها، چه ماتریس‌هایی و با چه ویژگی‌هایی می‌توانند خطای بوجود آمده را تصحیح کنند. مقاله‌های [۶۱] و [۴] مراجع اصلی این فصل هستند.

یک آزمون گروهی دارای خاصیت d -قدرت تشخیص است اگر آن آزمون توانایی تشخیص d فرد بیمار را داشته باشد. در فصل سوم شرط لازم و کافی برای اینکه یک ماتریس متناظر با آزمون گروهی، دارای خاصیت d -قدرت تشخیص باشد را در قالب یک قضیه عنوان و اثبات می‌کنیم. نشان می‌دهیم که آزمون گروهی ایستا برای شناسایی حداکثر دو

بیمار از میان 2^t نفر نیاز به حداقل $2t$ آزمون دارد.

در بخش سوم از این فصل از گراف‌ها برای ساختار آزمون گروهی ایستا برای $d = 2$ استفاده می‌کنیم. نشان می‌دهیم که اگر M ماتریس وقوع گراف G باشد آنگاه M دارای خاصیت d -قدرت تشخیص است اگر و تنها اگر گراف G فاقد دور سه یا چهارتایی باشد. واکیل با استفاده از قضیه‌ی مطرح شده گرافی را طراحی نمود که با استفاده از آن می‌توان برای شناسایی حداکثر دو بیمار از میان ۱۷۵ نفر از یک آزمون گروهی بهینه با ۵۰ گروه استفاده کرد. مرجع این فصل [۳۱] است.

در فصل چهارم براساس مراجع [۶] و [۵]، الگوریتم‌های آزمون گروهی را برای مدل معمولی، مدل پوچگر، مدل هم‌تافت و مدل هم‌تافت با پوچگر ارائه می‌دهیم و همچنین هر یک از این الگوریتم‌ها را به‌حالت خطا آزاد تعمیم می‌دهیم. در انتهای این فصل آزمون گروهی معمولی را به مدل آستانه تعمیم می‌دهیم و دو الگوریتم متناظر با حالت خاص مدل آستانه و مدل عمومی مدل آستانه را معرفی می‌کنیم. در حالت خاص مدل آستانه یک ارتباط به‌ظاهر نامربوط بین این حالت و مسئله‌ی جستجوی گراف پیدا می‌شود.

در فصل پنجم مسأله آزمون گروهی با پوچگرها را با دو حالت بررسی می‌کنیم. در حالت اول تعداد دقیق افراد بیمار را می‌دانیم و در حالت دوم کران بالا از افراد بیمار را داریم. همچنین در این فصل ساختارهای ترکیبیاتی را ارائه می‌دهیم که با استفاده از آن‌ها کران پایین بهینه روی تعداد تست‌هایی که برای شناسایی افراد بیمار در آزمون گروهی با پوچگرها لازم است را بدست می‌آوریم. این نتایج براساس مقاله‌های [۸] و [۱۶] هستند.

آنچه که در فصل ششم مورد مطالعه قرار می‌گیرد اینست که با فرض معلوم بودن d ، برای چه مقادیری از n در آزمون گروهی ایستا، آزمون انفرادی بهینه است. فرض می‌کنیم که برای ثابت d ، $N(d)$ بزرگترین مقدار n باشد که برای آن آزمون انفرادی بهینه است. در این فصل اثبات می‌کنیم که با در نظر گرفتن شرایطی در الگوریتم ایستا، $N(d) = (d+1)^2$ و همچنین اثبات می‌کنیم که برای $d = 1, 2, 3, 4$ ، $N(d) = (d+1)^2$ است. نتایج این فصل براساس مرجع [۲۰] است.

فهرست مطالب

ز	لیست جداول
ژ	لیست تصاویر
۱	۱ مقدمه و تاریخچه
۱	۱.۱ تاریخچه
۳	۲.۱ آزمون گروهی ترکیبیاتی
۵	۳.۱ آزمون گروهی ایستا
۱۰	۴.۱ آزمون گروهی و کاربردهای آن در ملکول شناسی
۱۳	۱.۴.۱ مدل خطا آزاد در آزمون گروهی
۱۵	۲ ماتریس‌های d - مجزا، d - جداسدنی و $\bar{d}$ - جداسدنی
۱۵	۱.۲ مقدمه
۱۶	۲.۲ حالت خطا آزاد
۲۰	۳.۲ قضایا و نتایج مهم
۲۶	۳ کران‌های بهینه برای آزمون گروهی
۲۶	۱.۳ مقدمه
۲۶	۲.۳ ساختار آزمون گروهی ایستا و خاصیت d - قدرت تشخیص
۳۰	۳.۳ ساختار آزمون گروهی روی گراف پترسن تعمیم یافته
۴۰	۴.۳ آزمون گروهی با چند مرحله
۴۳	۴ الگوریتم‌های آزمون گروهی ایستا و مدل آستانه
۴۳	۱.۴ مقدمه
۴۳	۲.۴ الگوریتم‌های مدل معمولی
۴۵	۱.۲.۴ الگوریتم مدل خطا آزاد
۴۷	۳.۴ مدل خطا آزاد با پوچگر (EI)
۴۹	۱.۳.۴ مدل معمولی خطا آزاد با پوچگر (GEI)
۵۰	۴.۴ الگوریتم‌های کدگشا برای مدل هم‌تافت خطا آزاد (EC)
۵۱	۱.۴.۴ الگوریتم مدل هم‌تافت خطا آزاد با پوچگر (EIC)

۵۴	الگوریتم‌های ایستا برای آزمون گروهی آستانه	۵.۴
۵۴	حالت بدون اختلاف ($g = 0$) مدل آستانه	۱.۵.۴
۵۸	حالت عمومی مدل آستانه	۲.۵.۴
۶۱	۵ ساختارهای ترکیبیاتی جدید و کاربردهای آن برای آزمون گروهی با پوچگرها	
۶۱	مقدمه	۱.۵
۶۲	تعداد دقیق افراد بیمار را نمی‌دانیم	۲.۵
۶۲	الگوریتم‌هایی با یک مرحله	۱.۲.۵
۶۷	کدهای ابر پوشش دهنده‌ی تعمیم یافته و انتخاب‌گرها	۳.۵
۷۳	یک الگوریتم با دو گام برای مسأله آزمون گروهی با پوچگر	۴.۵
۷۵	ماتریس‌های شمول	۵.۵
۸۱	۶ چه موقع برای آزمون گروهی ایستا، آزمون انفرادی بهینه است؟	
۸۱	مقدمه	۱.۶
۸۲	یک شرط لازم	۲.۶
۸۴	شرط لازم و کافی تحت $\bar{\lambda} = 1$	۳.۶
۸۹	بدست آوردن $N(d)$ برای مقادیر کوچک d	۴.۶
۱۰۶	مراجع	
۱۱۳	واژه‌نامه فارسی به انگلیسی	

لیست جداول

۲۸	۱.۳ نمایش $(X, f(X))$ برای هر X که $w(X) \leq d$
۷۶	۱.۵ $\delta(۴, ۲, ۳)$
۷۷	۲.۵ $\pi(۲, ۴, ۲, ۳)$

لیست تصاویر

۳۴		گراف پترسن	۱.۳
۳۴		نمایشی دیگر از گراف پترسن	۲.۳
۳۵		نمایشی از گراف پترسن تعمیم یافته	۳.۳
۳۶		گراف فاقد دور سه و چهارتایی	۴.۳
۳۸		گراف پیشنهادی واکیل	۵.۳

فصل ۱

مقدمه و تاریخچه

۱.۱ تاریخچه

برخلاف بسیاری از مسائل ریاضی که به قرن‌ها پیش و منابع مختلف برمی‌گردند منشأ **آزمون گروهی**^۱ تا حد زیادی به حادثه‌ی نسبتاً اخیر جنگ جهانی دوم (۴۳ – ۱۹۴۲) پیوندخورده و به روبرت دورفمن^۲ نسبت داده می‌شود [۱۳]. میلیون‌ها سرباز در جنگ حضور داشتند که از بین آن‌ها فقط تعداد اندکی، ۱۰۰۰ مورد، مبتلا به ویروس سیفلیس شده بودند. دورفمن و دیوید روزن‌بلات^۳ در یکی از شعبه‌های تحقیقاتی اداره‌ی آمار مالی واشنگتن کار می‌کردند، در جلسه‌ای که در این اداره با حضور تعدادی اقتصاددان برگزار شده بود تمام حضار به این نتیجه رسیدند که اگر برای پیدا کردن تعداد اندکی بیمار، نمونه خون میلیون‌ها سرباز را به‌صورت انفرادی مورد تجزیه و تحلیل آزمایشگاهی قرار دهیم عمل بی‌فایده‌ای است. آزمون گروهی توسط یکی از افراد در جلسه مطرح شد ولی از طرف دیگر فردی اظهار داشت که ممکن است گروه بندی کردن نمونه‌های خون مشکلات اقتصادی داشته باشد و این ایده رد شد. اما دورفمن و همکارش روزن‌بلات (دورفمن در مقالاتش به همکاری روزن‌بلات اشاره می‌کند) این ایده را به‌صورت جدی پذیرفتند. بعد از مدت کوتاهی دورفمن یک یادداشت چهار صفحه‌ای را در جلسه‌ی انجمن آمار واشنگتن ارائه نمود که در مجله سالیانه‌ی آمار ریاضی منتشر شد.

با توجه به گزارش اصلی [۹] از دورفمن:

^۱group testing

^۲Robert Dorfman

^۳ David Rosenblatt

هر سربازی که در معرض آزمایش خون میکروب سیفیلیس قرار می‌گیرد آزمایش به دو مرحله تقسیم می‌شود:

۱. نمونه خون افراد دریافت می‌شود.

۲. نمونه خون در معرض یک تجزیه و تحلیل آزمایشگاهی قرار می‌گیرد که وجود یا عدم وجود آنتی ژن سیفیلیس را نشان می‌دهد.

وقتی این فرآیند استفاده شود n تجزیه و تحلیل آزمایشگاهی به منظور شناسایی افراد مبتلا شده به ویروس از میان جمعیتی به اندازه‌ی n نیاز است.

با توجه به روش پیشنهاد شده میکروب به صورت زیر شناسایی می‌شود.

مقداری از نمونه خون دریافت شده از افراد را به گروه‌های پنج تایی تقسیم می‌کنیم، نمونه خون‌های هر گروهی را با هم ترکیب کرده و در یک مخزن می‌ریزیم حال نمونه خون داخل مخزن را در معرض تجزیه و تحلیل آزمایشگاهی قرار می‌دهیم اگر هیچ یک از پنج نمونه خون داخل مخزن حاوی میکروب سیفیلیس نباشد جواب آزمایش منفی، ولی اگر لااقل یکی از نمونه خون‌های داخل مخزن حاوی این میکروب باشد جواب آزمایش مثبت است. آزمایش‌های منفی کنار گذاشته شده ولی آزمایش‌های مثبت باید به صورت انفرادی مجدداً بررسی شوند.

در آن روز این ایده‌ی امیدبخش، گروه‌بندی کردن نمونه‌های خون، به صورت کاربردی مطرح نشده بود ولی امروزه ایده‌ی دورفمن تأثیرات جدید و مهمی بر روی جامعه‌ی پزشکی و صنعت گذاشته است.

مسئله آزمایش خون دورفمن در سال ۱۹۵۰ باعث تدوین کتاب معروف فلر^۴ در زمینه احتمالات تحت عنوان مقدمه‌ای بر تئوری احتمال و کاربرد آن، شد [۱۸].

سوبل^۵ و گرول^۶ [۳۲] دو دانشمند از آزمایشگاه بل در مقاله‌های خود به مفهوم جدیدی از آزمون گروهی اشاره نمودند برای مثال می‌توان به آزمون گروهی در صنعت اشاره کرد. دورفمن، سوبل و گرول آزمون گروهی را تحت مدل احتمالی بررسی نموده‌اند برای مثال یک توزیع احتمال بر روی تمام نواقص ضمیمه کردند و هدف را به حداقل

^۴ Feller

^۵ Sobel

^۶ Groll

رساندن تعداد آزمایش‌های لازم برای شناسایی نواقض قرار دادند.

آزمون گروهی ترکیبیاتی ابتدا توسط لی^۷ مورد مطالعه قرار گرفت [۲۵].

در این پایان نامه به آزمون گروهی ترکیبیاتی پرداخته می‌شود.

۲۰۱ آزمون گروهی ترکیبیاتی

گروهی از افراد بیمار و سالم را در نظر می‌گیریم به‌طوری‌که تعداد افراد بیمار نسبت به تعداد کل افراد گروه عدد خیلی کوچکی باشد. دقت کنید که با توجه به داده‌های تجربی در یک جامعه می‌توانیم نسبت افراد بیمار به کل جامعه را یک داده ثابت فرض کنیم.

هدف شناسایی افراد بیمار است و برای این کار آزمایش خون از تمامی افراد گروه گرفته می‌شود. حال اگر قرار باشد نمونه خون‌های گرفته شده از افراد گروه به‌صورت انفرادی مورد آزمایش قرار بگیرند ممکن است به دلایل ذیل مشکلاتی را به‌دنبال داشته باشند:

الف. بعضی از آزمایش‌ها وجود دارند که تجزیه‌ی آن‌ها هزینه‌های بالایی را دربردارد.

ب. تجزیه و تحلیل بعضی از آزمایش‌ها زمان زیادی را سپری می‌کند و همچنین کمبود دستگاه‌های آزمایشگاهی گرفتن آزمون انفرادی در زمان کوتاه را غیرممکن می‌سازد.

لذا یک ابزار مهم برای شناسایی بیماران آزمایش گروهی است.

فرض کنید یک گروه n -نفری داریم و می‌خواهیم افراد بیمار را از بین آن‌ها شناسایی کنیم. کل افراد این گروه را با N نمایش می‌دهیم. هر زیرمجموعه‌ی ناتهی از N مانند G را یک گروه می‌نامیم. نتیجه آزمایش روی یک گروه منفی است اگر همه اعضای گروه سالم باشند و مثبت است اگر لااقل یک فرد بیمار در G قرار داشته باشد.

یک آزمون گروهی شامل t گروه $G_1, G_2, \dots, G_t$ است به‌طوری‌که نتیجه آزمایش این گروه‌ها همه افراد بیمار را مشخص می‌کند.

^۷ Li

در آزمون گروهی هدف این است که با فرض وجود حداکثر d بیمار در میان جمعیتی به اندازه‌ی مشخص n با حداقل آزمایش ممکن بتوانیم افراد بیمار را شناسایی کنیم. که در نظر گرفتن دو نکته زیر در کاهش تعداد آزمایش‌ها اهمیت دارد:

الف. از آنجایی که تعداد بیماران بطورنسی خیلی کوچکتر از تعداد کل افراد است اگر اندازه گروه‌ها بزرگ باشد آنگاه برای بدست آوردن نتیجه مثبت نیاز به یک آزمون خیلی حساس داریم لذا اندازه گروه‌ها در عمل دارای محدودیت است.

ب. با توجه به نوع آزمایش برای هر شخص تعداد گروه‌های شامل این شخص دارای محدودیت است. برای مثال مقدار نمونه خون دریافتی از هر شخص محدود است و لذا تعداد گروه‌هایی که می‌توانند از این نمونه خون استفاده کنند نمی‌تواند زیاد باشد.

بطور کلی الگوریتم‌های آزمون گروهی به دو دسته تقسیم می‌شوند:

۱. الگوریتم‌های پویا.^۸

۲. الگوریتم‌های ایستا.^۹

یک الگوریتم آزمون گروهی را **پویا** می‌نامیم هرگاه آزمایش روی گروه‌ها به ترتیب (یکی یکی) انجام شود و در انتخاب یک گروه برای آزمایش نتیجه آزمایش گروه‌های قبلی مشخص هستند. یک الگوریتم آزمون گروهی را **ایستا** می‌نامیم اگر تمام گروه‌ها بطور همزمان مورد آزمایش قرار بگیرند و بعد از انجام تمام آزمایش‌ها بتوانیم افراد سالم و غیر سالم را شناسایی کنیم، در این الگوریتم مسأله اصلی چگونه گروه‌بندی کردن افراد است.

از آنجا که داشتن اطلاعات اضافی کمک می‌کند تا گروه‌بندی بهتری داشته باشیم لذا یک الگوریتم پویا به تعداد

^۸ sequential algorithms

^۹ nonadaptive algorithms

گروه‌های کمتری برای آزمایش نیاز دارد، و از آنجا که الگوریتم‌های ایستا بطور همزمان انجام می‌شوند لذا در زمان صرفه‌جویی می‌شود. بین الگوریتم‌های پویا و ایستا الگوریتم‌های چندمرحله‌ای^{۱۰} وجود دارد. در این پایان نامه به بررسی الگوریتم‌های ایستا می‌پردازیم.

۳.۱ آزمون گروهی ایستا

به یک آزمون گروهی ایستا می‌توان یک $(0, 1)$ -ماتریس، $M = (m_{ij})$ نسبت داد به‌طوری‌که هر سطر آن متناظر با یک گروه و هر ستون آن متناظر با یک شخص است. شخص j -ام به گروه i -ام تعلق دارد اگر و تنها اگر $m_{ij} = 1$ و در غیر این صورت برابر صفر است.

همچنین یک ستون C از M را می‌توان به عنوان یک زیرمجموعه از سطرهاى M در نظر گرفت که اعضای این زیرمجموعه دقیقاً همان سطرهایی از M هستند که درایه‌ی متناظر آن‌ها در C برابر یک است. S را زیرمجموعه‌ای از ستون‌های ماتریس M فرض می‌کنیم، $U(S)$ را اجتماع زیرمجموعه‌ی S از ستون‌های ماتریس M در نظر می‌گیریم $U(S)$ جمع دودویی ستون‌های S است که:

$$1 + 1 = 1 \quad 1 + 0 = 1 \quad 0 + 0 = 0.$$

برای مثال در ماتریس زیر فرض می‌کنیم $S = \{c_1, c_2, c_4\}$:

c_1	c_2	c_3	c_4	$U(S)$
۱	۰	۰	۰	۱
۰	۱	۰	۱	۱
۰	۱	۰	۱	۱
۰	۰	۱	۰	۰
۰	۰	۱	۰	۰

تعداد درایه‌های یک واقع در یک ستون برابر وزن یک ستون است در ماتریس فوق وزن ستون c_3 برابر دو است. بطور مشابه می‌توان یک سطر را به‌وسیله‌ی مجموعه‌ای از ستون‌های برخوردکننده به آن سطر که دارای درایه‌ی یک است نمایش دهیم و وزن یک سطر را تعداد درایه‌های یک آن سطر در نظر می‌گیریم.

^{۱۰} multi stage

برای دو بردار دودویی $V = (v_1, v_2, \dots, v_t)$ و $V' = (v'_1, v'_2, \dots, v'_t)$ است اگر و تنها اگر برای هر $1 \leq i \leq t$ ، $v_i \leq v'_i$ باشد.

فرض کنید یک گروه n -نفری داریم که در بین آن‌ها حداکثر d بیمار وجود دارد مجموعه‌ی p -عضوی P را مجموعه‌ی تمام بیماران در یک گروه n -نفری در نظر می‌گیریم که اندازه‌ی آن حداکثر برابر d است.

الف. بردار دودویی $D_{n \times 1}$ بردار اشخاص بیمار^{۱۱} نامیده می‌شود که برای $j = 1, 2, \dots, n$ ، اگر درایه‌ی $D_j = 1$ یعنی شخص j -ام بیمار است و در غیر این صورت $D_j = 0$.

ب. بردار دودویی $R_{t \times 1}$ بردار نتیجه^{۱۲} نامیده می‌شود. برای $i = 1, 2, \dots, t$ ، اگر درایه‌ی $R_i = 1$ یعنی نتیجه i -امین گروه مثبت است و در غیر این صورت $R_i = 0$.

برای ماتریس داده شده M و بردار D بردار نتیجه R به صورت زیر بدست می‌آید:

$$R_{t \times 1} = M_{t \times n} D_{n \times 1}.$$

به عبارت دیگر MD ترکیب خطی ستون‌های M است. مثلاً اگر $C_1, C_2, \dots, C_p$ ستون‌های متناظر با افراد بیمار باشند:

$$R^t = C_1 \oplus C_2 \oplus \dots \oplus C_p$$

و به وضوح بردار نتیجه برابر است با اجتماع ستون‌های متناظر با افراد بیمار:

$$R = U(P).$$

^{۱۱} vector of defective items(vod)

^{۱۲} result vector

می‌خواهیم ماتریس‌هایی را ارائه دهیم که متناظر با یک الگوریتم ایستا هستند.

کاتر^{۱۳} و سینگلوتن^{۱۴} دو مفهوم d -مجزا^{۱۵} و $\bar{d}$ -جداشدنی^{۱۶} را برای $(0, 1)$ -ماتریس، $M_{t \times n}$ مطرح کردند [۲۴].

فرض کنید M یک ماتریس است که می‌تواند افراد بیمار را شناسایی کند واضح است که برای دو زیرمجموعه متمایز حداکثر d -عضوی S و S' از ستون‌های M باید $U(S) \neq U(S')$. به یک ماتریس با این خاصیت یک ماتریس $\bar{d}$ -جداشدنی می‌گوییم.

ماتریس M را d -جداشدنی^{۱۷} می‌گوییم هرگاه برای هر دو زیرمجموعه متفاوت d -عضوی S و S' از ستون‌های M داشته باشیم $U(S) \neq U(S')$.

اما بررسی داشتن چنین شرط برای ماتریس ممکن است کاری شلوغ بنظر برسد لذا برای کدگشایی آسان الگوریتم‌ها حالت ضعیف دیگری را جایگزین می‌کنیم.

در ماتریس M اگر هیچ ستونی در اجتماع d ستون دیگر از ماتریس نباشد آن‌گاه ماتریس M ، d -مجزا نامیده می‌شود یا به عبارت دیگر برای هر $1 + d$ ستون $C_0, C_1, \dots, C_d$ از M داشته باشیم:

$$|C_0 \setminus \bigcup_{i=1}^d C_i| \geq 1.$$

یعنی اینکه حداقل یک سطر موجود است که C_0 در آن سطر دارای درایه یک و هر ستون C_i ، $1 \leq i \leq d$ ، در آن سطر دارای درایه صفر است. این بدین معنی است که اگر $C_1, C_2, \dots, C_d$ افراد بیمار باشند و C_0 فرد سالم باشد آن‌گاه حداقل یک گروه وجود دارد که شامل فرد سالم است ولی شامل هیچ بیماری نیست لذا نتیجه این آزمون منفی است و بیماران شناسایی می‌شوند.

بنابراین اگر برای ستون C از ماتریس d -مجزای M داشته باشیم $C \leq U(P)$ آن‌گاه $C \in P$. زیرا برای هر ستون

^{۱۳} Kautz

^{۱۴} Singleton

^{۱۵} d -disjunct

^{۱۶} $\bar{d}$ -separable

^{۱۷} d -separable

$P \notin C$ حداقل یک سطر وجود دارد که C را از اجتماع ستون‌های P جدا می‌کند.

با استفاده از تعاریف می‌توان نتیجه گرفت که هر ماتریس که دارای خاصیت d -جداشدنی ($\bar{d}$ -جداشدنی، d -مجزا) باشد، آن‌گاه یک ماتریس k -جداشدنی ($\bar{k}$ -جداشدنی، k -مجزا)، $1 \leq k < d < n$ ، نیز است و همچنین روابط زیر را داریم [۱۳]:

$$1 + \overline{d} - \text{جداشدنی} \Leftarrow d - \text{مجزا} \Leftarrow \bar{d} - \text{جداشدنی} \Leftarrow d - \text{جداشدنی}$$

برای n نفر که در میان آن‌ها حداکثر d بیمار داریم $t(d, n)$ را کمترین تعداد آزمون‌هایی که برای شناسایی d بیمار مورد نیاز است، تعریف می‌کنیم و همچنین $n(d, t)$ را بزرگترین مقدار n که برای آن یک ماتریس، $d, M_{t \times n}$ -مجزا داریم، تعریف می‌کنیم.

اسپنسر^{۱۸} [۳۳] نشان داد که با t آزمون، بزرگترین مقدار n برای اینکه یک ماتریس ۱-مجزا داشته باشیم برابر است با $n(1, t) = \binom{t}{\lfloor \frac{t}{2} \rfloor}$.

همان‌طور که قبلاً اشاره شد در یک آزمون انفرادی برای n نفر نیاز به n آزمایش است حال با فرض معلوم بودن d سوال این است که:

برای چه مقدار از n آزمون انفرادی بهینه است؟

هویو،^{۱۹} وانگ^{۲۰} و وینگ^{۲۱} [۱۹] حدس زدند که آزمون انفرادی بهینه است اگر و تنها اگر $n \leq 3d$ ، آن‌ها شرط لازم را اثبات کرده و شرط کافی برای $n \leq 2.5d$ اثبات شد. دو^{۲۲} و وانگ [۱۰] شرط کافی را به $n \leq 2.625d$ توسعه دادند و نتیجه اثبات شده توسط وانگ [۲۲] را که نشان داده بود حکم برای $n \leq 2d$ برقرار است، بهبود بخشیدند ولی حدس هنوز باز باقی مانده است. برای مقدار ثابت d ، $N(d)$ را بزرگترین عدد طبیعی n در نظر می‌گیریم که برای n نفر و d بیمار، آزمون انفرادی بهینه است. باسالایگو^{۲۳} [۱۴] کران پایینی را

^{۱۸} Spencer

^{۱۹} Hu

^{۲۰} Hwang

^{۲۱} Wang

^{۲۲} Du

^{۲۳} Bassalygo

برای $N(d)$ ارائه نمود.

$$N(d) \geq \binom{d+2}{2}. \quad \text{لم ۱.۳.۱}$$

اردیش^{۲۴}، فرینکال^{۲۵} و فوردی^{۲۶} [۵۰] حدس زدند که:

$$\lim_{d \rightarrow \infty} \frac{N(d)}{d^2} = 1$$

و در حالت قوی تر

$$N(d) \leq (d+1)^2.$$

آنها همچنین توضیح دادند که می توانند حالت قوی را برای $d \leq 3$ اثبات کنند.

در این پایان نامه اثبات خواهیم کرد برای زمانی که $d+1$ برابر توانی از یک عدد اول است، $N(d) \leq (d+1)^2$

است و لذا رابطه زیر را اثبات می کنیم:

$$\binom{d+2}{2} \leq N(d) \leq (d+1)^2.$$

^{۲۴}Erdős
^{۲۵}Frankl
^{۲۶}Füredi

۴.۱ آزمون گروهی و کاربردهای آن در ملکول شناسی

آزمون گروهی ابزاری است مهم که برای مسائل مختلفی مانند آزمایش خون، کانال‌های ارتباطی، تئوری کدگذاری و سایر مسائل می‌تواند بکار رود.

اخیراً آزمون گروهی در نظریه‌ی ملکول‌شناسی^{۲۷} مفید واقع شده است. برای مثال آزمایش‌های کلونی و دنباله‌ای با استفاده از اثر خنثی شدن^{۲۸} [۴۰، ۵۴].

امروزه محققان روشی به نام واکنش زنجیره‌ای^{۲۹} *PCR* را بکار می‌برند تا یک قطعه انتخاب شده از *DNA* را تکثیر کنند (کلون کردن *DNA*). واکنش‌گر به کمک این واکنش می‌تواند از یک نمونه در عرض چند ساعت چند میلیون نسخه از *DNA* مورد نظر بسازد.

در این واکنش ابتدا نمونه‌ای از *DNA* دو رشته‌ای را که می‌خواهند آن را تکثیر کنند حرارت می‌دهند حرارت باعث جدا شدن دو رشته‌ی *DNA* از هم می‌شود، البته اگر این دو رشته مجدداً در کنار هم قرار بگیرند تمایل به درهم تنیده شدن را دارند این تمایل به درهم تنیده شدن رشته‌ها را خنثی شدن می‌نامند، این مخلوط را سپس سرد می‌کنند و یک قطعه‌ی کوچک از *DNA* به نام پرایمر را به آن اضافه می‌کنند، پرایمر را به صورت مصنوعی می‌سازند، پرایمر به نقاطی از *DNA* که همانندسازی از آن‌ها آغاز می‌شود متصل می‌شود سپس آنزیم *DNA* و نوکلئوتیدهای آزاد را به مخلوط اضافه می‌کنند تا همانند سازی آغاز شود آنزیم *DNA* شروع می‌کند به اتصال نوکلئوتیدهای مکمل به پرایمر و باعث طویل شدن رشته *DNA* می‌شود نتیجه‌ی این فرایند تولید دو ملکول *DNA* است که با یکدیگر و با ملکول اولیه مشابه هستند فرایند گرمادهی و همانندسازی بارها و بارها تکرار می‌شود و هرچند دقیقه یکبار نمونه *DNA* دو برابر می‌شود و در زمان کوتاهی میلیون‌ها نسخه از نمونه‌ی اولیه بدست می‌آید [۱۲].

در مسأله‌ی آزمون کلونی با یک آزمایش مشخص می‌شود که کدامیک از کلون‌ها به صورت ناجور به هم تنیده شده‌اند اگر کلونی با جنسی ناجور پیوندخورده باشد آن کلون مثبت است و در غیر این صورت منفی است.

^{۲۷}molecular biology

^{۲۸} hybridization

^{۲۹}polymer chain resaction

نوع دیگری از کلون‌ها بنام **پوچگرها**^{۳۰} وجود دارد که کلون‌های مثبت را خنثی می‌کنند یعنی اینکه وجود یک پوچگر در هر گروه صرفنظر از وجود کلون‌های مثبت نتیجه منفی می‌دهد. مدل پوچگر ابتدا توسط فَرِیک^{۳۱} [۵۱] در سال ۱۹۹۷ مطرح شد و سپس مورد مطالعه‌ی دِ بونیز^{۳۲} [۴۴] در سال ۲۰۰۵، دِ بونیز و وُکرو^{۳۳} [۴۵] در سال ۱۹۹۸، چانگ^{۳۴} [۴۲] در سال ۲۰۰۷، وانگ و لایو^{۳۵} [۲۳] در سال ۲۰۰۳ قرار گرفت.

حال فرض می‌کنیم افراد همان ملکول‌ها هستند، در برخی از کاربردها بجای اینکه خاصیت مثبت و یا منفی بودن روی یک ملکول تعریف شود روی زیرمجموعه‌ای از ملکول‌ها تعریف می‌شود چنین مدلی را مدل **هم‌تافت**^{۳۶} می‌نامیم که در ابتدا توسط تورنای^{۳۷} [۳۴] در سال ۱۹۹۹ مطرح شد. مدل آزمون گروهی هم‌تافت به سایر مسائل مانند آزمون گراف، توزیع کلید امن^{۳۸} و سایر مسائل مرتبط است [۳۹، ۵۲، ۴۹، ۲، ۴۳].

در مدل هم‌تافت خانواده $P = \{P_i\}$ از زیرمجموعه‌های N را در نظر می‌گیریم به‌طوری‌که در کنار هم قرار گرفتن زیرمجموعه‌ای از ملکول‌ها (P_i) منجر به ایجاد یک پیامد بیولوژی و نتیجه‌ی مثبت می‌شود. مجموعه‌ای از ملکول‌ها هم‌تافت نامیده می‌شوند و به اعضای مجموعه‌ی P هم‌تافت‌های مثبت گفته می‌شود. هدف این است که با حداقل آزمایش‌ها بتوانیم هم‌تافت‌های مجموعه‌ی P را شناسایی کنیم.

به‌عنوان نکته‌ای مهم در مدل هم‌تافت فرض اصلی بر این است که هیچ دو هم‌تافت، X و X' در رابطه‌ی $X' \subseteq X$ صدق نمی‌کنند زیرا اگر هم‌تافتی مانند X شامل هم‌تافت بیمار X^+ باشد آن‌گاه در آزمون گروهی، در همه‌ی گروه‌هایی که هم‌تافت X وجود دارد هم‌تافت X^+ نیز ظاهر می‌شود و بنابراین نتیجه‌ی همه‌ی گروه‌هایی که X ظاهر می‌شود مثبت است. لذا هم‌تافت X نمی‌تواند شناسایی شود.

^{۳۰}inhibitor^{۳۱}Farach^{۳۲}De Bonis^{۳۳}Vaccaro^{۳۴}Chang^{۳۵}Liu^{۳۶}complex^{۳۷}Torney^{۳۸} secure key distribution

همانطور که اشاره شد مدل هم‌تافت می‌تواند به آزمون گروهی روی ابرگراف H ارتباط داده شود به‌طوری‌که هر ملکول به‌عنوان یک رأس و هر هم‌تافت نشان‌دهنده‌ی یک یال از ابرگراف H است و در آزمون گراف هدف این است که زیرگراف پنهان P ، که شامل مجموعه‌ای از یال‌های مثبت است را شناسایی کنیم. فرض می‌کنیم که گراف H یک گراف از مرتبه r باشد (هر یال شامل حداکثر r رأس باشد) و تنها چیزی که می‌دانیم این است که $|P| \leq d$. برای $(0, 1)$ - ماتریس M و S یک زیرمجموعه از ستون‌های آن $\cap S$ را فصل مشترک همه‌ی ستون‌های واقع در S تعریف می‌کنیم.

فرض کنید H یک ابرگراف با n رأس است. $(0, 1)$ - ماتریس M را که دارای n ستون است $(H_{\bar{r}} : d)$ - مجزا می‌گوییم اگر به ازای هر $d + 1$ یال $e_0, e_1, \dots, e_d$ از گراف $H_{\bar{r}}$ داشته باشیم:

$$\cap e_0 \not\subseteq \cup_{i=1}^d (\cap e_i).$$

با استفاده از یک ماتریس که دارای خاصیت فوق است می‌توان زیرگراف P را شناسایی کرد. زیرا برای هر یالی که خارج از P در نظر بگیریم حداقل یک گروه وجود دارد که شامل آن یال است ولی هیچ یک از یال‌های زیرگراف P را دربر ندارد و لذا نتیجه آزمون این گروه منفی است.

یک $(0, 1)$ - ماتریس $(d, r]$ - مجزا نامیده می‌شود اگر برای هر $d + r$ ستون $C_1, C_2, \dots, C_{d+r}$ از آن داشته باشیم:

$$|\cap_{i=1}^r C_i \setminus \cup_{i=r+1}^{d+r} C_i| \geq 1.$$

یعنی اینکه به‌ازای هر $d + r$ ستون حداقل یک سطر موجود است که تمام r ستون در این سطر دارای درایه یک و هر یک از d ستون دیگر دارای درایه‌ی صفر هستند.

ماتریس $(d, r]$ - مجزا ابتدا توسط افراد میچل^{۳۹} و پای پر^{۴۰} [۵۳] مطرح شد. بعدها چنین ماتریسی تحت نام‌های

^{۳۹}Mitchell

^{۴۰}Piper

خانواده پوشش آزاد ^{۴۱} و کدهای ابر پوشش دهنده ^{۴۲} برای کاربرد در زمینه‌ی توزیع کلید امن مورد مطالعه قرار گرفت [۵۸، ۵۹].

چن ^{۴۳} [۴۳] یک ارتباط بین مدل هم‌تافت و مسأله‌ی توزیع کلید امن برقرار کرد و نشان داد:

$$(d, r] - \text{مجزایی} \iff (H_{\bar{r}} : d) - \text{مجزایی برای هر } H_{\bar{r}}.$$

لذا برقراری چنین رابطه‌ای نتیجه می‌دهد که هر ماتریس $(d, r]$ - مجزا می‌تواند مسأله‌ی مدل هم‌تافت را حل کند.

۱.۴.۱ مدل خطا آزاد در آزمون گروهی

در کاربردهای ملکول شناسی از خطاهای آزمون نمی‌توان صرف‌نظر کرد در صورت وجود خطا، نتایج آزمایش بر روی گروه‌ها غیر قابل اعتماد بوده و بسیاری از نتایج مثبت و منفی اشتباه می‌شود بنابراین نیاز به ارائه‌ی تعریفی با توانایی شناسایی بیماران با فرض وجود خطا مطلوب است.

ماتریس دودویی $M_{t \times n}$ ، (d, z) - مجزا نامیده می‌شود اگر برای هر $d + 1$ ستون $C_0, C_1, \dots, C_d$ داشته باشیم:

$$|C_0 \setminus \bigcup_{i=1}^d C_i| \geq z.$$

یعنی اینکه حداقل z سطر وجود دارند که در آن سطرها C_0 دارای درایه یک و هر ستون C_i ، $1 \leq i \leq d$ ، دارای درایه صفر است.

وقتی $z = 1$ تعریف فوق با تعریف d - مجزا بودن بودن یکسان است [۱۱]، در این حالت بجای $(d, 1)$ - مجزا، d - مجزا می‌گوییم.

نشان خواهیم داد که با فرض وجود حداکثر $\lfloor \frac{z-1}{p} \rfloor$ خطا در بردار خروجی یک ماتریس (d, z) - مجزا می‌تواند حداکثر d بیمار را شناسایی کند.

^{۴۱}cover - free families

^{۴۲}superimposed codes

^{۴۳}Chen

(۱, ۰) - ماتریس $M_{t \times n}$ ، $(d, r, z]$ - مجزا نامیده می‌شود اگر برای هر $d + r$ ستون داشته باشیم:

$$|\bigcap_{i=1}^r C_i \setminus \bigcup_{i=r+1}^{d+r} C_i| \geq z.$$

یعنی اینکه حداقل z سطر وجود دارد که در این سطرها تمام r ستون دارای درایه یک و هر یک از d ستون دیگر دارای درایه‌ی صفر هستند.

فرض می‌کنیم $t(n, d, r, z)$ کوچکترین t ای باشد که برای آن یک ماتریس $M_{t \times n}$ ، $(d, r, z]$ - مجزا داریم.

چن، فو^{۴۴} و وانگ کران بالای زیر را برای $t(n, d, r, z)$ ارائه دادند [۱]:

$$t(n, d, r, z) < z \left(\frac{k}{r}\right)^r \left(\frac{k}{d}\right)^d [1 + k(1 + \ln(\frac{n}{k} + 1))]. \quad (۱.۱)$$

که $k = d + r$.

فصل ۲

ماتریس‌های d - مجزا ، d - جداشدنی و $\bar{d}$ - جداشدنی

۱.۲ مقدمه

سه نوع از $(0, 1)$ - ماتریس‌ها وجود دارند که در مسأله آزمون گروهی مورد مطالعه قرار گرفته‌اند. ماتریس d - جداشدنی توسط اردیش و موسر^۱ [۱۷] برای $d = 2$ مورد مطالعه قرار گرفت، ماتریس $\bar{d}$ - جداشدنی توسط کاتز و سینگلوتن [۲۴] برای یک نوع خاص از کدها به نام UD_d بررسی شد و همچنین ماتریس d - مجزا نیز توسط همین اشخاص مورد مطالعه قرار گرفت. اخیراً چن و وانگ [۳] روشی برای تبدیل یک ماتریس d - جداشدنی به یک ماتریس d - مجزا ارائه دادند. دو^۲ و وانگ [۱۲] این نتایج را به مدل خطا آزاد گسترش دادند.

^۱Moser

^۲Du

۲.۲ حالت خطا آزاد

ماتریس‌های d -جداشدنی و $\bar{d}$ -جداشدنی، مانند ماتریس d -مجزا، دارای حالت خطا آزاد هستند یعنی اینکه با وجود خطا در خروجی می‌توانند افراد بیمار را شناسایی کنند.

(۰, ۱)-ماتریس M ، (d, z) -**جداشدنی** نامیده می‌شود اگر برای هر دو زیرمجموعه‌ی d -عضوی S و S' از ستون‌های M داشته باشیم:

$$|U(S) \Delta U(S')| \geq z.$$

اگر هر دو زیرمجموعه‌ی d -عضوی در بالا را به دو زیرمجموعه‌ای با حداکثر d عضو تغییر دهیم، آن‌گاه ماتریس M ، $(\bar{d}, z)$ -**جداشدنی** نامیده می‌شود.

همان‌طور که در فصل یک اشاره شد ماتریس M ، (d, z) -مجزا نامیده می‌شود اگر برای هر زیرمجموعه‌ی d -تایی S از ستون‌های M و هر ستون دیگری مانند C داشته باشیم:

$$|C \setminus U(S)| \geq z.$$

یعنی اینکه حداقل z سطر موجود است که ستون C در آن سطرها دارای درایه یک است و هر ستون S در آن سطرها دارای درایه صفر است.

دو بردار دودویی X و Y به طول n را در نظر می‌گیریم **فاصله همینگ**^۳ X و Y را که با $H(X, Y)$ نمایش می‌دهیم به صورت زیر تعریف می‌شود:

$$H(X, Y) = |\{i; 1 \leq i \leq n, X(i) \neq Y(i)\}|.$$

^۳Hamming distance

گزاره ۱.۲.۲. [۱۶] فرض می‌کنیم $1 \leq d \leq n$ و M یک ماتریس $t \times n$ ، (d, z) -مجزا باشد. اگر $H(x, y)$ فاصله همینگ بین دو بردار دودویی به طول t باشد و همچنین S و T زیرمجموعه‌های متمایز از ستون‌های M به اندازه حداکثر d باشند آن‌گاه:

الف. اگر $S \subset T$ آن‌گاه $H(U(S), U(T)) \geq z$

ب. اگر $S \not\subset T$ و $T \not\subset S$ سپس $H(U(S), U(T)) \geq 2z$

اثبات. الف. از آنجایی که $S \subset T$ لذا حداقل یک ستون $C_0 \in T$ که $C_0 \notin S$ وجود دارد. لذا بنا به خاصیت (d, z) -مجزا بودن حداقل z سطر موجود است که C_0 در آن سطرها دارای درایه‌ی یک و تمام ستون‌های S در این سطرها برابر صفر هستند که حکم برقرار است.

ب. بنابر فرض مسأله ستون $C_0 \in T$ که $C_0 \notin S$ و ستون $C_1 \in S$ که $C_1 \notin T$ موجود است که بنا به خاصیت (d, z) -مجزا بودن مجموعه‌های (C_0, S) و (C_1, T) به وسیله حداقل z سطر از M مجزا هستند و لذا نتیجه برقرار است.

□

با استفاده از مطالب قبلی و گزاره (۱.۲.۲)، یک ماتریس با داشتن خاصیت (d, z) -مجزایی می‌تواند برای تصحیح کردن یک تعداد خطا استفاده شود.

فرض می‌کنیم اندازه P حداکثر برابر d باشد سپس داشتن خاصیت (d, z) -مجزایی می‌تواند حداکثر $\lfloor \frac{z-1}{4} \rfloor$ خطا را تصحیح کند. فرض می‌کنیم حداکثر $\lfloor \frac{z-1}{4} \rfloor$ خطا در بردار نتیجه R اتفاق بیفتد و ε بردار حاصل از خطای مشاهده شده باشد لذا $H(R, \varepsilon) = H(U(P), \varepsilon) \leq \lfloor \frac{z-1}{4} \rfloor$. به وسیله بررسی کردن همه زیرمجموعه‌های S از ستون‌های M به اندازه حداکثر d برای فقط یک S داریم $H(U(S), \varepsilon) \leq \lfloor \frac{z-1}{4} \rfloor$ که $S = P$ است. زیرا اگر $S \neq P$ آن‌گاه بنابر گزاره (۱.۲.۲) قسمت (الف) داریم:

$$H(U(S), \varepsilon) + H(\varepsilon, R) \geq H(U(S), R) = H(U(S), U(P)) \geq z$$

و از آنجا که $H(\varepsilon, R) \leq \lfloor \frac{z-1}{2} \rfloor$ لذا $H(U(S), \varepsilon) > \lfloor \frac{z-1}{2} \rfloor$.

حال اگر P دارای اندازه‌ی دقیقاً برابر d باشد سپس یک ماتریس (d, z) -مجزا می‌تواند $z - 1$ خطا را تصحیح کند. فرض می‌کنیم حداکثر $z - 1$ خطا در بردار خروجی R اتفاق بیفتد که بردار حاصل از خطای مشاهده شده را δ می‌نامیم در این صورت داریم $H(R, \delta) = H(U(P), \delta) \leq z - 1$ به وسیله بررسی کردن همه زیرمجموعه‌های S از ستون‌های M به اندازه d برای فقط یک S داریم $H(U(S), \delta) \leq z - 1$ که $S = P$ است. زیرا بنا به گزاره (۱.۲.۲) قسمت (ب) اگر $S \neq P$ آن‌گاه:

$$H(U(S), \delta) + H(\delta, R) \geq H(U(S), R) = H(U(S), U(P)) \geq 2z$$

و از آنجایی که $H(\delta, R) \leq z - 1$ لذا $H(U(S), \delta) > z$ که این تناقض است.

حال سوالی که پیش می‌آید این است که آیا با فرض حداکثر d عضوی بودن P ، یک ماتریس (d, z) -مجزا می‌تواند $z - 1$ خطا را تصحیح کند؟ جواب این سوال هم بله و هم خیر است. جواب خیر است اگر آزمون گروهی ایستا مطلوب باشد.

مثال نقض زیر این موضوع را نشان می‌دهد.

ماتریس M_* ، (d, z) -مجزا با $d = 2$ و $z = 2$ است. فرض می‌کنیم دو ستون اول ماتریس متناظر با بیماران باشند در این صورت بردار نتیجه و بدون خطای R_1 را داریم. حال اگر مجموعه‌ی بیماران یک عضوی باشد فرض می‌کنیم ستون اول ماتریس متناظر با آن بیمار باشد که در این صورت بردار نتیجه و بدون خطای R_2 را داریم. حال اگر بردار خطا δ باشد آن‌گاه فاصله همینگ دو بردار R_1 و R_2 از δ برابر یک خواهد بود. بنابراین با یک خطا نمی‌توان زیرمجموعه‌ی مثبت و منحصر به فرد را شناسایی کرد.

$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad R_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \delta = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

مثال فوق نشان می‌دهد که با خاصیت ایستایی و استفاده از یک ماتریس (d, z) - مجزا به‌طوری که اندازه P حداکثر d باشد نمی‌توان $z - 1$ خطا را تصحیح کرد. اما با استفاده از یک روش بدیهی دومرحله‌ای ماتریس (d, z) - مجزا می‌تواند $z - 1$ خطا را تصحیح کند. در روش بدیهی دومرحله‌ای، مرحله دوم اضافه می‌شود که در این مرحله تمامی افراد یک گروه نسبتاً کوچک بطور انفرادی مورد آزمایش قرار می‌گیرند و لذا هر تست انفرادی می‌تواند خطا را تصحیح کند. حال این روش را در زیر توصیف می‌کنیم.

فرض می‌کنیم زیرمجموعه‌ی مثبت P دارای اندازه‌ی حداکثر d باشد و همچنین فرض می‌کنیم حداکثر $z - 1$ خطا در بردار نتیجه R اتفاق بیفتد و بردار مشاهده شده‌ی خروجی δ باشد. زیرمجموعه‌های S از ستون‌های M به اندازه حداکثر d با $H(U(S), \delta) \leq z - 1$ را در نظر می‌گیریم. فرض می‌کنیم S_1 و S_2 دو تا از چنین زیرمجموعه‌هایی باشند، در این صورت یا $S_1 \subset S_2$ یا $S_2 \subset S_1$ که اگر چنین نباشد بنا بر گزاره (۱.۲.۲) قسمت (ب) $H(U(S_1), U(S_2)) \geq 2z$ است. لذا یا $H(U(S_1), \delta) > z - 1$ یا $H(U(S_2), \delta) > z - 1$ که تناقض است. لذا مجموعه‌ی تمام زیرمجموعه‌هایی از ستون‌های M به اندازه حداکثر d مانند S که $H(U(S), \delta) \leq z - 1$ یک زنجیر هستند و از آنجایی که $H(U(P), \delta) = H(R, \delta) \leq z - 1$ پس P یکی از اعضای زنجیر است که با آزمون انفرادی مشخص می‌شود.

۳.۲ قضایا و نتایج مهم

لم ۱.۳.۲. [۱۲] اگر ماتریس M ، (d, z) -مجزا باشد آنگاه یک ماتریس $(d, 2z)$ -جداشدنی است.

اثبات. فرض می‌کنیم M یک ماتریس (d, z) -مجزا است که S و S' دو زیرمجموعه‌ی d -تایی متمایز و دلخواه از ستون‌های ماتریس M باشند. سپس هر $C \in S \setminus S'$ دارای حداقل z سطر است که در $U(S')$ نیست و همچنین هر ستون $C' \in S' \setminus S$ دارای حداقل z سطر است که در $U(S)$ نیست بنابراین $U(S)$ و $U(S')$ در حداقل $2z$ سطر متمایز هستند که این معادل است با اینکه ماتریس M ، $(d, 2z)$ -جداشدنی است. $\square$

لم ۲.۳.۲. [۴] اگر M یک ماتریس d -جداشدنی با n ستون باشد که $d < n$ ، آن‌گاه M برای هر عدد مثبت k ، $k \leq d$ -جداشدنی است.

اثبات. فرض می‌کنیم ماتریس M ، d -جداشدنی باشد اما برای k ، $k < d$ -جداشدنی نباشد یعنی دو زیرمجموعه‌ی k -تایی S و S' از ستون‌های ماتریس M وجود دارد به‌طوری‌که $U(S) = U(S')$. فرض می‌کنیم C_x ستونی در همسایگی S یا S' باشد لذا $C_x \cup U(S) = C_x \cup U(S')$. $d - k$ ستون از چنین C_x هایی را به هر دو زیرمجموعه‌ی S و S' اضافه می‌کنیم که این امر منجر به ایجاد دو زیرمجموعه‌ی متمایز S_d و S'_d با d ستون می‌شود که $U(S_d) = U(S'_d)$. بنابراین M ، d -جداشدنی نیست.

حال اگر فقط $l < d - k$ تا از چنین ستون‌های C_x موجود باشد در این صورت $d - k - l$ جفت از ستون‌های (C_y, C_z) انتخاب می‌کنیم به‌طوری‌که C_y در S وجود دارد ولی در S' نیست و C_z در S' است اما در S نیست. سپس $C_z \cup U(S) = U(S') = U(S) = C_y \cup U(S')$. بنابراین این زوج‌ها می‌توانند جایگزین بقیه‌ی C_x ها باشند. لذا $d - k$ ستون C_x و (C_y, C_z) همواره وجود دارند به‌طوری‌که منجر به ایجاد دو زیرمجموعه‌ی متمایز S_d و S'_d با d ستون می‌شوند که $U(S_d) = U(S'_d)$ و این تناقض است. $\square$

توجه شود شرط $d < n$ در لم (۲.۳.۲) یک شرط لازم است. مثال زیر این موضوع را نشان می‌دهد.

ماتریس M_1 بطوربدیهی ۳-جداشدنی است ولی ۲-جداشدنی نیست.

$$M_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

حال لم (۲.۳.۲) را به حالت خطا آزاد زیر تعمیم می‌دهیم.

لم ۳.۳.۲. [۴] اگر ماتریس M با n ستون (d, z) -جداشدنی ($d < n$) باشد آن‌گاه برای هر عدد مثبت $k \leq d$ ، (k, z) -جداشدنی است.

اثبات. کافی است تا نشان دهیم ماتریس M ، $(d-1, z)$ -جداشدنی است. فرض می‌کنیم M ، $(d-1, z)$ -جداشدنی نباشد لذا زیرمجموعه‌های متمایز S و S' شامل $d-1$ ستون از M موجود است که $|U(S) \Delta U(S')| < z$. حال اگر $|S \setminus S'| = |S' \setminus S| \geq 2$ آن‌گاه باید یک جفت ستون‌های (C_x, C_y) وجود داشته باشد به‌طوری‌که $C_x \in S \setminus S'$ و $C_y \in S' \setminus S$ از طرفی داریم:

$$|U(S \cup \{C_y\}) \Delta U(S' \cup \{C_x\})| \leq |U(S \cup \{C_y\}) \Delta U(S')| \leq |U(S) \Delta U(S')| < z$$

و این تناقض با خاصیت (d, z) -جداشدنی بودن ماتریس M است.

حال اگر $|S \setminus S'| = |S' \setminus S| = 1$ به‌وضوح $|S \cup S'| = d$ و از آنجایی که $d < n$ لذا یک ستون C از ماتریس M بجز ستون‌های S یا S' وجود دارد و از طرفی داریم:

$$|U(S \cup \{C\}) \Delta U(S' \cup \{C\})| \leq |U(S) \Delta U(S')| < z$$

و این نامساوی تناقض با (d, z) -جداشدنی بودن ماتریس M است.

لذا برای هر دو زیرمجموعه‌ی متمایز S و S' به اندازه $d-1$ از ستون‌های M داریم $|U(S) \Delta U(S')| \leq z$. پس

M ، $(d-1, z)$ -جداشدنی است و با ادامه دادن روند فوق می‌توان حکم قضیه را نتیجه گرفت. $\square$

قضیه ۴.۳.۲. [۴] اگر M یک ماتریس $(2d, z)$ -جداشدنی با n ستون که $n \geq 2d+1$ باشد آن گاه با اضافه کردن حداکثر $\lceil \frac{z}{4} \rceil$ سطر به ماتریس M می‌توان یک ماتریس $(d, \lceil \frac{z}{4} \rceil)$ -مجزا بدست آورد.

اثبات. فرض می‌کنیم M ، $(d, \lceil \frac{z}{4} \rceil)$ -مجزا نباشد لذا ستون C و زیرمجموعه‌ی S با d ستون چنان وجود دارد که $|C \setminus U(S)| < \lceil \frac{z}{4} \rceil$. حداکثر $\lceil \frac{z}{4} \rceil$ سطر را به ماتریس M طوری اضافه می‌کنیم که در این سطرها، ستون C دارای درایه‌ی یک باشد و ستون‌های S در این سطرها همگی برابر صفر باشند، که در این صورت داریم $|C \setminus U(S)| \geq \lceil \frac{z}{4} \rceil$. البته ممکن است یک زوج دیگری مانند (C', S') که C' ستونی از M و S' مجموعه‌ای از d ستون دیگر از M باشد وجود داشته باشد که $|C' \setminus U(S')| < \lceil \frac{z}{4} \rceil$ سپس می‌توان به روش بالا عمل کرد. ولی باید نشان دهیم که این فرایند خودمتناقض نیست یعنی نمی‌تواند دو جفت (C, S) و (C', S') به‌طوری که $C \in S'$ ، $|C \setminus U(S)| < \lceil \frac{z}{4} \rceil$ و $|C' \setminus U(S')| < \lceil \frac{z}{4} \rceil$ وجود داشته باشند.

فرض می‌کنیم دو جفت (C, S) و (C', S') که $|S| = |S'| = d$ ، توصیف شده به‌صورت بالا وجود داشته باشد. تعریف می‌کنیم:

$$S_* = \{C'\} \cup S \cup S', \quad S_1 = S_* \setminus \{C\}, \quad S_2 = S_* \setminus \{C'\}.$$

فرض می‌کنیم $|S_*| = s$ لذا:

$$s \leq 2d+1, \quad |S_1| = |S_2| = s-1 \leq 2d.$$

حال نشان می‌دهیم که تفاضل متقارن $U(S_1)$ و $U(S_2)$ کمتر از z است که این تناقض با $(2d, z)$ -جداشدنی بودن M است. از آنجایی که تنها ستونی که در S_1 است اما در S_2 نیست C' است و $|C' \setminus U(S')| < \lceil \frac{z}{4} \rceil$ داریم:

$$|U(S_1) \setminus U(S_2)| < \lceil \frac{z}{4} \rceil$$

و بطور مشابه داریم:

$$|U(S_2) \setminus U(S_1)| < \lceil \frac{z}{4} \rceil.$$

لذا $z < |U(S_1) \Delta U(S_2)|$ و این یعنی ماتریس M ، $(s-1, z)$ -جداشدنی نیست و این تناقض دارد با لم (۳.۳.۲) و لذا اثبات تمام است. $\square$

نتیجه ۵.۳.۲. [۴] اگر M یک ماتریس $2d$ -جداشدنی با n ستون که $n \geq 2d+1$ باشد آن گاه با اضافه کردن حداکثر یک سطر به M یک ماتریس d -مجزا بدست می آید.

اثبات. با قرار دادن $z = 1$ در قضیه ی (۴.۳.۲) حکم برقرار است. $\square$

مثال زیر نشان می دهد که شرط $n \geq 2d+1$ در نتیجه (۵.۳.۲) لازم است. فرض می کنیم:

$$M_2 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

به وضوح M_2 ، ۴-جداشدنی است که با اضافه کردن یک سطر جدید ۲-مجزا نخواهد بود.

بطور مشابه هر ماتریسی با $2d$ ستون $(2d, z)$ -جداشدنی است و با اضافه کردن $\lceil \frac{z}{2} \rceil$ سطر نمی توان یک ماتریس $(d, \lceil \frac{z}{2} \rceil)$ -مجزا با $2d$ ستون ساخت. به عنوان یک مثال M_3 یک ماتریس $(4, 4)$ -جداشدنی است اما با اضافه کردن ۲ سطر نمی توان یک ماتریس $(2, 2)$ -مجزا ساخت حتی $(1, 2)$ -مجزا هم نیست.

$$M_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

فرض می کنیم C_1, C_2, C_3 و C_4 چهار ستون از M_3 باشند، بنابر خاصیت $(1, 2)$ -مجزایی برای هر زوج مرتب (C_i, C_j) ، $i \neq j$ ، از ستون های ماتریس M_3 نیاز به حداقل دو سطر است که هر یک از این دو سطر شامل C_i باشد ولی شامل C_j نباشد اولین سطر توسط ماتریس M_3 فراهم می شود ولی برای دومین سطر نیاز به اضافه کردن سطرهای جدید به ماتریس است. برای زوج مرتب (C_i, C_j) ، $12 = 4 \times 3$ انتخاب وجود دارد لذا نیاز به دوازده جفت $(1, 0)$ است. یک سطر جدید حداکثر چهار زوج $(1, 0)$ را می تواند ایجاد کند، که این کار با قرار دادن دو درایه یک و دو درایه صفر انجام می شود، لذا دو سطر جدید حداکثر هشت زوج $(1, 0)$ از دوازده زوج را ایجاد می کند بنابراین اضافه کردن دو سطر برای داشتن خاصیت $(1, 2)$ -مجزایی کافی نیست.

نتیجه ۶.۳.۲. [۴] اگر ماتریس M ، $(2d, z)$ -جداشدنی با n ستون که $n \geq 2d + 1$ باشد آن گاه برای هر عدد مثبت $k \leq \lceil \frac{z}{4} \rceil$ می‌توان یک ماتریس (d, k) -مجزا با اضافه کردن حداکثر k سطر به M ساخت.

اثبات. در اثبات قضیه (۴.۳.۲) نشان دادیم که دو جفت (C, S) و (C', S') که $|C \setminus U(S)| < \lceil \frac{z}{4} \rceil$ و $C \in S'$ و $|C' \setminus U(S')| < \lceil \frac{z}{4} \rceil$ وجود ندارند. در حقیقت می‌توان $\lceil \frac{z}{4} \rceil$ را با یک عدد مثبتی مانند k جایگزین کرد، به طوری که تفاضل متقارن $U(S_1)$ و $U(S_2)$ کمتر از z است بنابراین برای هر $k \leq \lceil \frac{z}{4} \rceil$ می‌توانیم ماتریس (d, k) -مجزا را با اضافه کردن حداکثر k سطر به M ساخت. $\square$

لم ۷.۳.۲. [۴] ماتریس M ، $(\bar{d}, z)$ -جداشدنی است اگر و تنها اگر M ، (d, z) -جداشدنی و $(d-1, z)$ -مجزا باشد.

اثبات. فرض می‌کنیم M ، $(\bar{d}, z)$ -جداشدنی باشد اما $(d-1, z)$ -مجزا نباشد. لذا زیرمجموعه‌ی $(d-1)$ -عضوی S و ستون C از ستون‌های M وجود دارند که $|C \setminus U(S)| < z$ ، داریم:

$$|U(S \cup \{C\}) \Delta U(S)| = |U(S \cup \{C\}) \setminus U(S)| < z,$$

که این تناقض با $(\bar{d}, z)$ -جداشدنی بودن ماتریس M دارد. بنابراین ماتریس M ، $(d-1, z)$ -مجزا و به وضوح (d, z) -جداشدنی است.

فرض می‌کنیم M ، (d, z) -جداشدنی و $(d-1, z)$ -مجزا باشد کافی است نشان دهیم که برای هر دو زیرمجموعه‌ی حداکثر d -عضوی، X و Y از M داریم $|U(X) \Delta U(Y)| \geq z$.

اگر $|X| = |Y| \leq d$ ، با توجه به اینکه ماتریس M ، (d, z) -جداشدنی است و بنا بر لم (۳.۳.۲) داریم:

$$|U(X) \Delta U(Y)| \geq z.$$

حال فرض می‌کنیم $|X| < |Y| \leq d$ لذا ستون $C_y \in Y$ وجود دارد که متعلق به X نیست با استفاده از خاصیت $(d-1, z)$ -مجزایی داریم $|C_y \setminus U(X)| \geq z$ بنابراین $|U(X) \Delta U(Y)| \geq z$ لذا اثبات کامل می‌شود. $\square$

قضیه ۸.۳.۲. [۶۱] اگر ماتریس M ، d -جداشدنی باشد آن گاه برای هر $1 \leq k \leq d-1$ ماتریس M ، $\overline{k+1}$ -جداشدنی است اگر و تنها اگر M یک ماتریس k -مجزا باشد.

اثبات. شرط کافی: فرض می‌کنیم دو زیرمجموعه‌ی حداکثر $(k+1)$ -عضوی S و S' از ستون‌های M وجود داشته باشد به‌قسمی که $U(S) = U(S')$. حال با داشتن خاصیت d -جداشدنی، جمع هر دو زیرمجموعه‌ی متمایز و d -تایی از ستون‌های ماتریس M با هم برابر نیست، فرض می‌کنیم $|S| < |S'| \leq k+1$ بنابراین ستون C وجود دارد که $C \in S' \setminus S$ لذا داریم $C \leq U(S')$ و چون $U(S) = U(S')$ پس $C \leq U(S)$ حال اگر فرض کنیم $|S| = k$ با خاصیت k -مجزایی M به تناقض می‌رسیم.

شرط لازم: فرض می‌کنیم M ، k -مجزا نباشد لذا ستون C و مجموعه‌ی k -تایی S از ستون‌های ماتریس M وجود دارند که $C \leq U(S)$ حال با فرض $S' = S \cup \{C\}$ داریم $U(S) = U(S')$ که $|S|, |S'| \leq k+1$ بنابراین M ، $\overline{k+1}$ -جداشدنی نخواهد بود که این تناقض است. $\square$

قضیه ۹.۳.۲. [۶۱] اگر M یک ماتریس $2d$ -جداشدنی باشد آن گاه با اضافه کردن حداکثر یک سطر به ماتریس M می‌توان یک ماتریس $\overline{d+1}$ -جداشدنی ساخت.

اثبات. فرض می‌کنیم M ، $2d$ -جداشدنی باشد با توجه به قضیه (۵.۳.۲) با اضافه کردن حداکثر یک سطر به M می‌توان یک ماتریس d -مجزا بدست آورد و با توجه به قضیه (۸.۳.۲) چون $1 \leq d \leq 2d-1$ لذا M یک ماتریس $\overline{d+1}$ -جداشدنی است. $\square$

فصل ۳

کران‌های بهینه برای آزمون گروهی

۱.۳ مقدمه

سها^۱ و سین‌ها^۲ اولین اشخاصی بودند که مشاهده کردند برای $n \leq 5$ تعداد آزمون‌های گروهی با تعداد کل افراد برابر است [۵۶]. ساختار آزمون گروهی پویا و ایستا در کاهش تعداد آزمون‌های مورد نیاز برای شناسایی افراد بیمار از میان مجموعه‌ای با حداقل شش نفر مفید هستند. در این فصل کران بهینه‌ای را روی تعداد آزمون‌های مورد نیاز در آزمون گروهی پویا و ایستا بدست می‌آوریم و نشان می‌دهیم که آزمون گروهی ایستا برای شناسایی حداکثر دو بیمار از میان 2^t نفر نیاز به حداقل $2t$ آزمون دارد و آزمون گروهی چندمرحله‌ای پویا با ماکزیمم $2t$ آزمون گروهی را معرفی می‌کنیم و همچنین نشان می‌دهیم که تحت شرایطی روی اندازه گروه می‌توان با استفاده از گراف پترسن تعمیم‌یافته آزمون گروهی بهینه‌ای را ساخت.

۲.۳ ساختار آزمون گروهی ایستا و خاصیت d - قدرت تشخیص

تعریف ۱.۲.۳. [۴۱، ۵۵] یک آزمون گروهی دارای خاصیت d - قدرت تشخیص است اگر آن آزمون توانایی تشخیص d فرد بیمار را داشته باشد.

اگر یک آزمون گروهی دارای خاصیت d - قدرت تشخیص باشد آن‌گاه بردار نتیجه R در تناظر یک‌به‌یک با

^۱Saha

^۲Sinha

مجموعه افراد بیمار D است.

برای ۱ - قدرت تشخیص بودن باید بردارهای ستونی M متمایز از هم باشند. زیرا اگر $C_i = C_j$ که C_i ستون متناظر با شخص بیمار i و C_j ستون متناظر با شخص سالم j باشد در این صورت در هر گروهی که شخص j وجود دارد شخص i نیز هست لذا نتیجه هر آزمونی که شخص سالم j وجود دارد مثبت است که در این صورت نمی توان شخص j را شناسایی کرد.

برای ۲ - قدرت تشخیص بودن باید جمع دودویی هر دو ستون دلخواه با جمع دودویی دو ستون دلخواه دیگر متفاوت باشد. در غیر این صورت اگر جمع دودویی دو ستون متناظر با دو شخص بیمار a و b برابر جمع دودویی دو ستون متناظر با دو شخص سالم c و d باشد در این صورت در هر گروهی که c یا d و یا هر دو وجود دارد حداقل یک بیمار هم وجود دارد لذا گروه یا گروهایی را نمی توان یافت که نتیجه های آنها منفی باشند و شامل این دو فرد سالم باشند.

قضیه ۲.۲.۳. [۳۱] شرط لازم و کافی برای ماتریس $M_{t \times n}$ ، متناظر با یک آزمون گروهی، برای اینکه دارای خاصیت d - قدرت تشخیص باشد این است که به ازای هر دو بردار دودویی X و Y از مرتبه $1 \times n$ ، که $w(X) \leq d$ و $w(Y) \leq d$ داشته باشیم:

$$MX = MY \implies X = Y.$$

اثبات. فرض می کنیم ماتریس $M_{t \times n}$ دارای خاصیت d - قدرت تشخیص باشد و همچنین فرض می کنیم افراد متناظر با سطرهای $X_{n \times 1}$ همان افراد بیمار باشند که $MX = R_1$ و $MY = R_2$. لذا اگر $R_1 = R_2$ بنا به خاصیت d - قدرت تشخیص بودن M و اینکه $w(X) \leq d$ و همچنین بردار نتیجه در تناظر یک به یک با افراد بیمار است لذا $X = Y$.

حال عکس قضیه را اثبات می کنیم فرض می کنیم تابع $f(X) = MX$ یک به یک باشد:

$$f : \{X \mid X_{n \times 1}, w(X) \leq d\} \longrightarrow \{Y \mid Y_{t \times 1}, f(X) = Y\}.$$

جدول ۱.۳: نمایش $(X, f(X))$ برای هر X که $w(X) \leq d$

X	X_1	X_2	$\dots$	X_k
$F(X)$	$F(X_1)$	$F(X_2)$	$\dots$	$F(X_k)$

برای هر X که $w(X) \leq d$ ، $(X, f(X))$ را در جدول (۱.۳) می‌نویسیم.

که:

$$k = 1 + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{d}.$$

فرض می‌کنیم جواب آزمون را مشخص کرده‌ایم و بردار نتیجه Y شده است اگر تعداد بیماران حداکثر d باشد لذا:

$$Y \in \{f(X_1), f(X_2), \dots, f(X_k)\}.$$

فرض می‌کنیم $Y = f(X_i)$ در این صورت X_i افراد بیمار را مشخص می‌کند.

اثبات دیگری برای قسمت دوم: ادعا می‌کنیم که اگر $d \leq l$ ستون بگونه‌ای باشد که هر تست مثبت دارای حداقل

یک درایه‌ی ناصفر در این l ستون باشد آن‌گاه این l ستون همگی بیمار هستند. فرض می‌کنیم $C_1, C_2, \dots, C_l$

چنین ستون‌هایی باشند، ولی ستون‌های متناظر با افراد بیمار $C'_1, C'_2, \dots, C'_p$ باشند. لذا هر سطری که مثبت شده

است باید لااقل در یکی از ستون‌های $1 \leq i \leq k$ ، C'_i دارای درایه‌ی یک باشد و همچنین بنا به فرض داریم:

$$\bigcup_{i=1}^l C_i = \bigcup_{i=1}^p C'_i$$

و از آنجایی که بردار نتیجه در تناظر یک‌به‌یک با افراد بیمار است این تناقض است. $\square$

قضیه زیر برای $d = 1$ برقرار است.

قضیه ۳.۲.۳. [۲۱] فرض می‌کنیم $2^{t-1} \leq n < 2^t$ و ماتریس $M_{t \times n}$ ماتریسی باشد که ستون‌های آن نمایش t

بیت دودویی عدد j ، $j = 1, 2, \dots, n$ ، باشد آن‌گاه ماتریس $M_{t \times n}$ به‌عنوان ماتریس متناظر با ساختار یک آزمون

گروهی است که n تعداد افراد و t تعداد تست‌ها است و قدرت تشخیص یک بیمار را دارد.

اثبات. برای عدد x نمایش t بیت دودویی آن به صورت زیر است:

$$x = a_{t-1}2^{t-1} + a_{t-2}2^{t-2} + \dots + a_12 + a_0, \quad a_i \in \{0, 1\} \quad \forall i = 0, 1, \dots, t-1.$$

که ماتریس M به صورت زیر حاصل می شود:

$$M = \begin{pmatrix} 0 & 0 & 0 & \vdots \\ 1 & 1 & 1 & \vdots \\ 0 & 0 & 0 & \vdots \\ 0 & 1 & 1 & \vdots \\ 1 & 0 & 1 & \vdots \end{pmatrix}_{t \times n}$$

از آنجایی که نمایش دودویی اعداد $j = 1, 2, \dots, n$ دوه دو متمایز است پس ستون های ماتریس M متمایز بوده و دارای خاصیت ۱ - قدرت تشخیص است. $\square$

نتیجه ۴.۲.۳. [۳۱] اگر در قضیه (۳.۲.۳) بدانیم که دقیقاً یک فرد بیمار وجود دارد با اضافه کردن یک ستون تهی به ماتریس می توانیم با $2^t \leq n \leq 2^{t-1}$ کار کنیم.

برای ماتریس، M ، در قضیه (۳.۲.۳) ماکزیمم اندازه گروه برابر $\lceil \frac{\bar{n}}{4} \rceil$ است. زیرا با توجه به ساختار ماتریس M هر سطر G_i ، $i = 0, 1, \dots, t-1$ ، نشان دهنده ضریب های 2^i در بسط دودویی اعداد $1, 2, \dots, n$ است. حال هدف این است $|G_i|$ را بدست آوریم. به ازای ستون های $1, 2, \dots, 2^i - 1$ دارای 2^i درایه ی صفر در سطر G_i هستیم و به ازای ستون های $2^i, 2^i + 1, \dots, 2^{i+1} - 1$ دارای 2^i درایه ی یک در این سطر هستیم و به همین ترتیب در 2^i ستون بعدی دارای درایه صفر خواهیم بود حال با فرض n ستون ماکزیمم اندازه G_i برابر $\lceil \frac{\bar{n}}{4} \rceil$ است. با فرض $n = 2^t - 1$ ماکزیمم وزن ستون های ماتریس M برابر t است.

قضیه ۵.۲.۳. [۳۱] در حالت $n = 2^t$ و $d = 2$ هر آزمون گروهی ایستا به حداقل $2t$ آزمون نیاز دارد.

اثبات. فرض می کنیم ماتریس $M_{g \times n}$ دارای خاصیت ۲ - قدرت تشخیص باشد که $g \leq 2t - 1$. برای بردار نتیجه $2^g, Y_{g \times 1}$ حالت متفاوت وجود دارد و چون $g \leq 2t - 1$ پس حداکثر 2^{2t-1} حالت، بردار نتیجه داریم. از طرفی حداکثر انتخاب های ممکن برای بردار اشخاص بیمار زمانی که حداکثر دو بیمار داریم به صورت زیر است:

$$1 + \binom{n}{1} + \binom{n}{2} = 1 + 2^{t-1} + 2^{2t-1} > 2^{2t-1}$$

و این یعنی بردار نتیجه در تناظر یک به یک با بردار افراد بیمار نیست. لذا M قدرت شناسایی برای دو بیمار

را ندارد و این تناقض است. $\square$

۳.۳ ساختار آزمون گروهی روی گراف پترسن تعمیم یافته

در این بخش هدف این است که از گراف‌ها برای ساختار آزمون گروهی ایستا برای $d = 2$ استفاده کنیم. دو و وانگ بطور مختصر به این موضوع پرداخته‌اند [۱۱]. گراف G زوج مرتب شده‌ای مانند (V, E) است که V مجموعه‌ی غیرتهی است که رأس‌های گراف G نامیده می‌شود و E مجموعه‌ای از یال‌ها به صورت $\{v_i, v_j\}$ است که v_i و v_j رأس‌های گراف هستند. دو رأس مجاورند اگر و تنها اگر یالی بین این دو رأس در گراف وجود داشته باشد. فرض می‌کنیم $v_1, v_2, \dots, v_n$ و $e_1, e_2, \dots, e_m$ به ترتیب رأس‌ها و یال‌های گراف G باشند. ماتریس وقوع $M(G)$ یک ماتریس از مرتبه $n \times m$ است که (i, j) -امین درایه از آن یک است اگر و تنها اگر یال e_j به رأس v_i برخورد کرده باشد در غیر این صورت صفر است. از آنجایی که هر یال فقط از دو رأس می‌گذرد لذا هر ستون دارای وزن دو است. اگر یال چندگانه نداشته باشیم، ستون‌های M متمایزاند. یک دور n -تایی یک دنباله از n رأس متمایز به صورت $a_1, a_2, \dots, a_n$ است به طوری که برای هر i ، $1 \leq i \leq n$ ، (a_i, a_{i+1}) یال‌های گراف G بوده و $a_1 = a_{n+1}$.

قضیه ۱.۳.۳ [۲۱] گراف فاقد یال چندگانه G را در نظر می‌گیریم و فرض می‌کنیم M ماتریس وقوع گراف G باشد. M دارای خاصیت ۲-قدرت تشخیص است اگر و تنها اگر گراف G فاقد دور سه یا چهارتایی باشد.

اثبات. ($\Rightarrow$) فرض می‌کنیم M ، ماتریس وقوع گراف G باشد که دارای خاصیت ۲-قدرت تشخیص است و فرض می‌کنیم G دارای دور سه یا چهارتایی باشد. می‌توان بردارهای دودویی و متمایز X و Y با وزن حداکثر دو مانند زیر یافت که $MX = MY$ و بنابر قضیه (۲.۲.۳) این تناقض با خاصیت ۲-قدرت تشخیص بودن M است پس

G دارای دور سه یا چهارتایی نیست.

اگر G دارای دور سه تایی باشد دارای ماتریس وقوع زیر است:

$$M = \begin{pmatrix} 1 & 0 & 1 & . & . & . \\ 1 & 1 & 0 & . & . & . \\ 0 & 1 & 1 & . & . & . \\ 0 & 0 & 0 & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{pmatrix}$$

دو بردار دودویی و متمایز X و Y با وزن حداکثر دو مانند زیر وجود دارند:

$$X = \begin{pmatrix} 1 \\ 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}$$

که :

$$MX = MY = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ . \\ . \\ . \end{pmatrix}.$$

اگر G دارای دور چهارتایی باشد دارای ماتریس وقوع زیر است:

$$M = \begin{pmatrix} 1 & 0 & 0 & 1 & . & . \\ 1 & 1 & 0 & 0 & . & . \\ 0 & 1 & 1 & 0 & . & . \\ 0 & 0 & 1 & 1 & . & . \\ 0 & 0 & 0 & 0 & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{pmatrix}$$

در این صورت دو بردار دودویی به صورت زیر موجود است :

$$X = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

که :

$$MX = MY = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

($\Rightarrow$) فرض می‌کنیم گراف G دارای دور سه یا چهارتایی نیست و همچنین فرض می‌کنیم ماتریس M دارای خاصیت

۲- قدرت تشخیص نباشد لذا بنا بر قضیه (۲.۲.۳) دو بردار دودویی متمایز X و Y با حداکثر وزن دو وجود دارند

که $MX = MY$.

از آنجایی که گراف G فاقد یال چندگانه است، هر کدام از بردارهای X و Y دارای وزن دو هستند. برای اثبات

این ادعا حالت‌های زیر را در نظر می‌گیریم:

حالت اول: $w(X) = w(Y) = 1$

فرض می‌کنیم بردار X در سطر i -ام دارای درایه‌ی یک باشد. از آنجا که هر ستون M دارای وزن دو است لذا بردار

ستونی MX دارای دو درایه‌ی یک است که در سطرها‌ی متناظر ستون i -ام از M که دارای درایه‌ی ناصفر هستند،

قرار دارند. فرض می‌کنیم بردار Y در سطر $j \neq i$ دارای درایه‌ی یک باشد. مشابه حالت قبل MY دارای دو

درایه‌ی یک متناظر با سطرها‌ی M است که دارای درایه‌ی ناصفر در ستون j هستند و چون $j \neq i$ و ستون‌های

M متمایز از هم هستند لذا $MX \neq MY$ که این تناقض است.

حالت دوم: $w(X) = 1$ و $w(Y) = 2$ یا بالعکس.

در این حالت به‌وضوح $w(MX) = 2$ و $w(MY) = 3$ یا $w(MY) = 4$ است. لذا $MX \neq MY$ که این

تناقض است.

حالت سوم: $w(X) = 2$ و $w(Y) = 2$

لذا حالت سوم برقرار است که در این صورت $MX (= MY)$ دارای وزن سه یا چهار است. فرض می‌کنیم بردار X در سطریهای i و j دارای درایه‌ی ناصفر باشد حال اگر ستون‌های i و j از ماتریس M یک اشتراک داشته باشند در این صورت بردار MX دارای وزن سه است که با جابه‌جایی سطر و ستون‌های لازم زیر ماتریس زیر را از M خواهیم داشت:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

حال یال‌هایی که با ستون‌های فوق معادل‌اند تشکیل دور سه تایی می‌دهند.

حال اگر ستون‌های i و j از ماتریس M اشتراکی نداشته باشند در این صورت بردار MX دارای وزن چهار است که با جابه‌جایی سطر و ستون‌های لازم زیر ماتریس زیر را از M خواهیم داشت:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

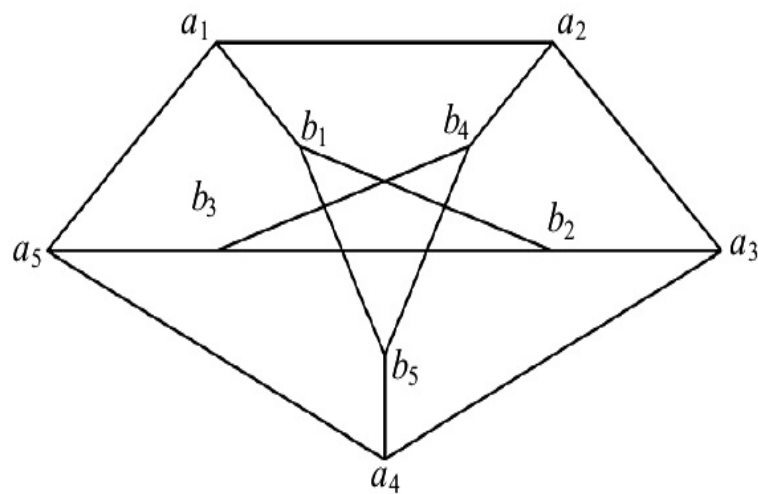
□ حال یال‌هایی که با ستون‌های فوق معادل‌اند تشکیل دور چهار تایی می‌دهند. لذا حکم برقرار است.

گراف‌هایی که در قضیه فوق صدق می‌کنند می‌توانند به گراف‌های ساختار آزمون گروهی ارجاع داده شوند. در ادامه گراف پترسن تعمیم یافته را معرفی می‌کنیم و نشان می‌دهیم که آن‌ها از نوع گراف‌های ساختار آزمون گروهی هستند. همچنین یک حل متناوب برای مسأله مطرح شده توسط واکیل^۳ را با استفاده از گراف پترسن تعمیم یافته ارائه می‌دهیم [۳۵، ۳۶]. گراف زیر، $P(10, 15)$ ، با ده رأس و پانزده یال **گراف پترسن**^۴ است.

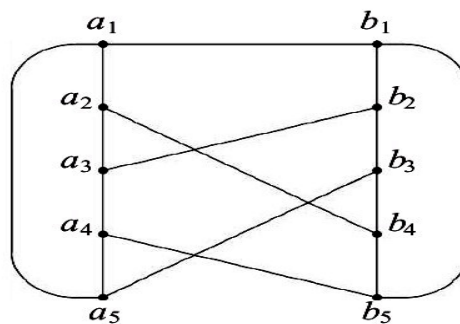
این گراف فاقد دور سه یا چهار تایی است و هر رأس از درجه سه است، که می‌توان آن‌را به صورت زیر نشان داد:

^۳Vakil

^۴Petersen graph



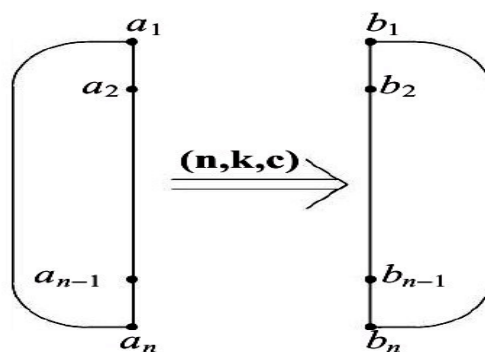
شکل ۱.۳: گراف پترسن



شکل ۲.۳: نمایشی دیگر از گراف پترسن

همچنین می‌توانیم گراف پترسن تعمیم یافته $P(2n, 3n)$ با $2n$ رأس و $3n$ یال برای $n \geq 5$ ، $n \neq 6$ به صورت زیر ساخت:

یک گراف با دو دور n -تایی مجزا مانند $A = (a_1, a_2, \dots, a_n)$ و $B = (b_1, b_2, \dots, b_n)$ را در نظر می‌گیریم. برای $i = 1, 2, \dots, n$ بین دو رأس a_i و b_j یال وجود دارد اگر $j \equiv ik + c \pmod{n}$ که $2 \leq k \leq n-2$ و ثابت k نسبت به n اول است و مقدار ثابت c به پیمانه n برده می‌شود لذا $1 \leq c \leq n-1$. مجموعه یال‌های اضافه



شکل ۳.۳: نمایشی از گراف پترسن تعمیم یافته

شده از A به B را با $A \xrightarrow{(n,k,c)} B$ نمایش می‌دهیم. برای مثال مجموعه یال‌های گراف پترسن در شکل (۲.۳) به صورت $B \xrightarrow{(5,3,3)} A$ است.

به راحتی می‌توان دید که گراف ایجاد شده به روش فوق بدون دور سه یا چهارتایی است.

فرض می‌کنیم a_i و b_j مجاور باشند لذا داریم:

$$j \equiv ik + c \pmod{n}.$$

چون A یک دور است لذا از a_i به a_{i+1} یال وجود دارد، حال اگر قرار باشد از a_{i+1} به b_j یالی وجود داشته باشد باید رابطه زیر برقرار باشد:

$$j \equiv (i+1)k + c \equiv ik + c + k \equiv j + k \pmod{n}.$$

چون k نسبت به n اول است لذا:

$$j + k \not\equiv j \pmod{n}.$$

در نتیجه گراف بدون دور سه‌تایی است.

چون A و B دور هستند لذا از a_i به a_{i+1} و از b_j به b_{j+1} یال وجود دارد. حال فرض می‌کنیم از a_i به b_j یالی وجود داشته باشد حال اگر از a_{i+1} به b_{j+1} یالی موجود باشد باید داشته باشیم:

$$j + 1 \equiv (i + 1)k + c \equiv ik + c + k \equiv j + k \pmod{n}.$$

چون $k \geq 2$ لذا

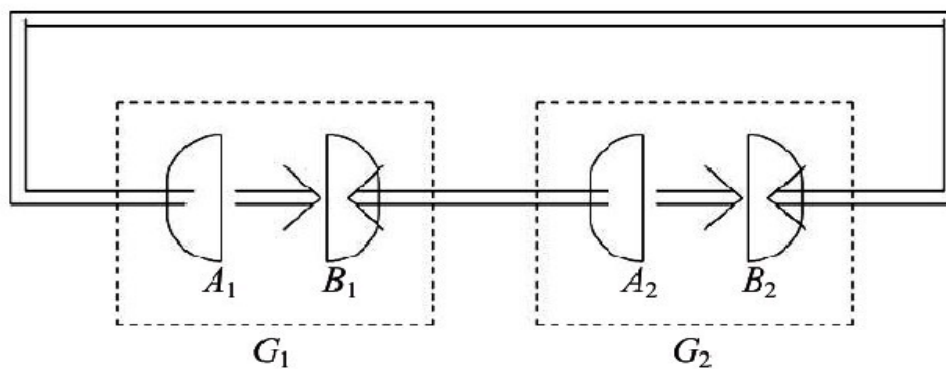
$$j + k \not\equiv j + 1 \pmod{n}.$$

دو گراف پترسن $P(2n, 3n)$ ، G_1 و G_2 با مجموعه یال‌هایی از A_1 به B_1 و از A_2 به B_2 را در نظر می‌گیریم که

$$A_2 \xrightarrow{(n,k,c)} B_2 \text{ و } A_1 \xrightarrow{(n,k,c)} B_1$$

دو گراف G_1 و G_2 را با اضافه کردن مجموعه یال‌هایی به صورت $A_2 \xrightarrow{(n,k,c)} B_1$ و همچنین $A_1 \xrightarrow{(n,k,c)} B_2$ که

$c \neq c_0$ ارتباط می‌دهیم.



شکل ۴.۳: گراف فاقد دور سه و چهارتایی

با انتخاب مقادیر ثابت k, c و c_0 می‌توان مانند بالا نشان داد که چنین گرافی دارای چرخه سه یا چهارتایی نیست.

فرایند اضافه کردن یال‌ها بین دوگراف G_1 و G_2 به وسیله $G_1 \xrightarrow{(n,k,c,c_0)} G_2$ نشان داده می‌شود به وضوح در گراف حاصل درجه هر رأس چهار است.

گراف پترسن، $G, p(2n, 3n)$ را در نظر می‌گیریم هر رأس G را با گراف پترسن $p(2m, 3m)$ و هر یال از آن را با مجموعه یال‌های (n, k, c, c_0) جایگزین می‌کنیم گراف G^* حاصل با $4mn$ رأس و $12mn$ یال است که فاقد دور سه یا چهارتایی است.

در واقع ابتدا با یک گراف آزمون گروهی (بدون دور سه و چهار تایی) شروع می‌کنیم و هر رأس را با یک گراف پترسن $p(2m, 3m)$ و هر یال را با مجموعه یال‌های (n, k, c, c_0) جایگزین می‌کنیم در این صورت یک گراف بدون دور سه و چهارتایی داریم لذا بنا به قضیه (۱.۳.۳) یک گراف آزمون گروهی حاصل می‌شود.

بنابراین در ساختار آزمون گروهی هر گرافی با حذف کردن زیرمجموعه‌ای از رأس‌ها و یال‌های برخوردکننده‌ی گراف‌های فوق حاصل می‌شود.

قواعدی که در ادامه آمده است می‌تواند در شناسایی افراد بیمار با استفاده از نتایج گراف آزمون گروهی مورد استفاده قرار گیرد.

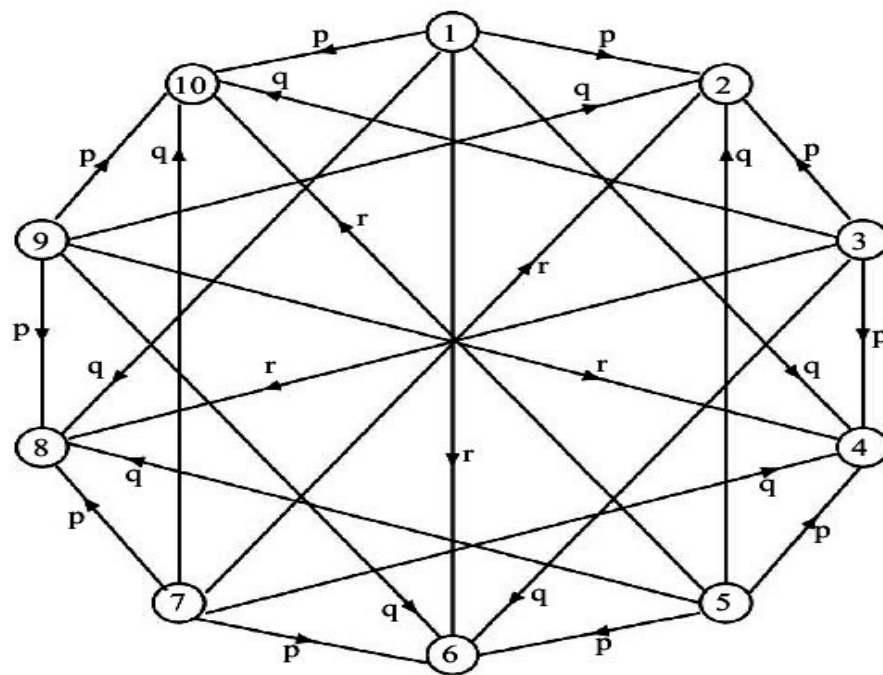
در گراف‌های آزمون گروهی یال‌ها نشان‌دهنده افراد و رأس‌ها نشان‌دهنده گروه‌ها هستند که یال‌های متناظر با افراد بیمار یال‌های بیمار و رأس‌های متناظر با گروه‌های مثبت رأس‌های مثبت نامیده می‌شوند. لذا هدف این است که یال‌های بیمار را براساس رأس‌های مثبت ارائه شده شناسایی کنیم. واضح است که نداشتن رأس مثبت دلیل به نداشتن یال مثبت است.

در مورد اینکه فقط دو رأس بیمار داریم یال ارتباط دهنده بین این دو رأس تنها یال بیمار است و این یعنی فقط یک فرد بیمار وجود دارد.

در مورد دو فرد بیمار سه یا چهار رأس بیمار داریم در مورد سه رأس بیمار دو یال ارتباط دهنده به این رأس‌ها یال

بیمار است (از آنجایی که در گراف دور سه تایی وجود ندارد بنابراین فقط دو یال ارتباط دهنده با این سه رأس وجود دارد). در مورد چهار رأس بیمار می‌تواند دو یا سه یال در زیر گرافی از این چهار رأس باشد (چون در گراف دور چهار تایی وجود ندارد) در مورد دو یال این دو یال به‌عنوان یال بیمار هستند و در مورد سه یال، این سه یال یک زنجیر به‌طول سه تشکیل می‌دهند که به‌وضوح دو یال انتهایی از این زنجیر یال‌های بیمار هستند. از آنجا که $d = 2$ نمی‌تواند بیشتر از چهار رأس بیمار باشد.

واکیل در موارد معینی ساختار تکنیک‌های بنیادی را برای ساختن آزمون گروهی بهینه استفاده می‌کند که برای $d = 2$ شامل یک آزمون گروهی بهینه با ۱۷۵ نفر و ۵۰ آزمون است.



شکل ۵.۳: گراف پیشنهادی واکیل

یک گراف شامل ده دور پنج‌تایی دوبه‌دو مجزای $C_1, C_2, \dots, C_{10}$ را در نظر می‌گیریم. در گراف فوق دایره i -ام دور پنج‌تایی C_i را برای $i = 1, 2, \dots, 10$ نمایش می‌دهد و پیکان‌هایی که آن‌ها را به هم وصل می‌کند مجموعه یال‌هایی است که به‌صورت زیر به گراف اضافه می‌شود.

برای $i = 1, 3, 5, 7$ مجموعه یال‌های اضافه شده از C_i به $C_{i\pm 1}$, $C_{i\pm 3}$ و C_{i+5} به صورت زیر است:

$$C_i \xrightarrow{p} C_{i\pm 1},$$

$$C_i \xrightarrow{q} C_{i\pm 3},$$

$$C_i \xrightarrow{r} C_{i+5}$$

که $p = (5, 2, 0)$, $q = (5, 2, 2)$ و $r = (5, 2, 1)$ به عنوان رابط‌ها ارجاع داده می‌شوند. به وضوح اگر رابط $x = (5, 2, c)$ از A به B وجود داشته باشد رابط B به A به صورت $\bar{x} = (5, 3, 2c)$ خواهد بود.

در گراف فوق دور سه‌تایی وجود ندارد. فرض می‌کنیم دور سه‌تایی C_a , C_b و C_c در گراف موجود باشد بدون از دست دادن کلیت مسأله فرض می‌کنیم a زوج باشد لذا b فرد و c زوج است و از آنجا که مجموعه یال‌هایی از دور زوج به زوج (یا فرد به فرد) وجود ندارد پس گراف بدون دور سه‌تایی است.

همچنین در گراف دور چهارتایی وجود ندارد زیرا چنین دوری در گراف باید دارای چهار دور پنج‌تایی C_a , C_b , C_c و C_d باشد فرض می‌کنیم a فرد است لذا c فرد و b, d زوج است. این دور چهارتایی باید یکی از رابط‌های زیر باشد:

$$q\bar{r}q\bar{r}, p\bar{q}q\bar{q}, p\bar{q}q\bar{r}, p\bar{q}r\bar{q}, p\bar{r}q\bar{q}, p\bar{p}q\bar{r}, p\bar{p}r\bar{q}, p\bar{q}\bar{p}r, p\bar{q}p\bar{q}, p\bar{r}p\bar{r}, p\bar{p}p\bar{q}$$

با بررسی این حالت‌ها روی گراف به سادگی می‌بینیم که هیچکدام منجر به ایجاد یک دور چهارتایی نخواهند شد.

۴.۳ آزمون گروهی با چند مرحله

در این بخش آزمون گروهی پویا با چند مرحله را در نظر می‌گیریم که در هر مرحله یک آزمون ایستا انجام می‌گیرد.

قضیه ۱.۴.۳ [۳۶] برای $d = 2$ و x^y نفر، یک آزمون گروهی پویا با دو مرحله وجود دارد که در مرحله اول xy آزمون و در مرحله دوم حداکثر $y - 1$ آزمون انجام می‌شود.

اثبات. در ابتدا x^y فرد را با اعداد 0 تا $x^y - 1$ شماره‌گذاری می‌کنیم و آن‌ها را با y رقم بر مبنای x نمایش می‌دهیم. برای $i = 1, 2, \dots, y$ و $j = 0, 1, \dots, x - 1$ گروه‌های G_i^j را تشکیل می‌دهیم که شامل تمامی افرادی است که عدد j را در مکان i -ام دارند. به‌وضوح برای ثابت j ، G_i^j یک گروه x^{y-1} نفری است لذا برای هر i ، $G_i^0, G_i^1, \dots, G_i^{x-1}$ یک بخش با x^y فرد را تشکیل می‌دهند.

$$i = 1 \quad G_1^0 \ G_1^1 \ \dots \ G_1^{x-1},$$

$$i = 2 \quad G_2^0 \ G_2^1 \ \dots \ G_2^{x-1},$$

$$\vdots \quad \quad \quad \vdots$$

$$i = y \quad G_y^0 \ G_y^1 \ \dots \ G_y^{x-1}.$$

در مرحله اول xy گروه G_i^j را آزمایش می‌کنیم. اگر نتیجه‌ی مثبتی نداشته باشیم آن‌گاه هیچ فرد بیماری نداریم. اگر به‌ازای هر i ، فقط یک مقدار j موجود باشد که G_i^j مثبت است آن‌گاه دقیقاً یک فرد بیمار داریم که با توجه به گروه‌های مثبت به‌راحتی می‌توان فرد بیمار را شناسایی کرد. از طرفی چون $d = 2$ لذا به‌ازای هر i برای یک یا دو عدد j ، G_i^j نتیجه مثبت دارد. اگر حداقل یک i موجود باشد که برای دو مقدار از j ، G_i^j دارای نتیجه مثبت باشد آن‌گاه قطعاً دو فرد بیمار موجود است و به سراغ مرحله دوم می‌رویم.

دیدیم که برای هر i حداکثر برای دو مقدار از j ، G_i^j دارای نتیجه مثبت است. حال اگر دو مقدار از j چنین باشد فرض می‌کنیم آن دو عنصر i_1 و i_2 باشند و اگر فقط برای یک مقدار از j ، G_i^j مثبت باشد i_1 و i_2 را برابر مقدار

j قرار می‌دهیم. بنابراین به‌ازای هر i دو عنصر و در کل $2y$ عنصر داریم (که نیاز به متمایز بودن نیست). بنابراین برای هر y - تایی حداکثر 2^y حالت داریم که از بین این حالات دو حالت وجود دارد که افراد متناظر با آن‌ها بیمار هستند. بدون اینکه به کلیت مسأله خللی وارد شود فرض می‌کنیم که برای $i = 1$ ، i_1 و i_2 دو مقدار متفاوت باشند. از میان 2^y نفر در بالا، گروه G را که شامل تمامی افرادیست که در مکان اول آن‌ها رقم 1_1 قرار گرفته است، در نظر می‌گیریم. بنابراین این گروه شامل 2^{y-1} نفر است که دقیقاً یک فرد بیمار دارد که بنابر نتیجه (۴.۲.۳) با استفاده از $1 - y$ تست می‌توان فرد بیمار را شناسایی کرد. بعد از تشخیص اولین فرد بیمار به‌راحتی فرد بیمار دوم شناسایی می‌شود. پس برای تشخیص حداکثر دو فرد بیمار از میان x^y نفر در آزمون پویا نیاز به $1 - y + xy$ آزمون است. $\square$

قضیه ۲.۴.۳. [۳۱] یک آزمون گروهی چندمرحله‌ای وجود دارد که می‌تواند حداکثر دو فرد بیمار را از میان 2^n نفر با حداکثر $2n$ آزمون شناسایی کند.

در ابتدا 2^n نفر را از $1 - 2^n$ مرتب می‌کنیم و آن‌ها را با n بیت دودویی نمایش می‌دهیم. به‌وضوح گروه‌های G_1^0 و G_1^1 هر کدام دارای 2^{n-1} نفراند لذا این دو گروه شامل تمام 2^n هستند. اگر نتیجه آزمایش روی گروه‌های G_1^0 و G_1^1 منفی باشد یعنی هیچ فرد بیماری وجود ندارد و اگر هر دو گروه، نتیجه‌ی مثبت داشته باشند آن‌گاه دقیقاً دو فرد بیمار وجود دارد. حال اگر نتیجه یک گروه مثل $G_1^{j_1}$ مثبت باشد آن‌گاه یک یا دو فرد بیمار وجود دارد. در این حالت دو گروه $G_1^{j_1}$ و $G_1^{j_2}$ را در نظر می‌گیریم به‌طوری‌که برای $j = 0, 1$ ، $G_1^{j_1}$ شامل تمامی افرادیست که j_1 به‌عنوان اولین بیت و j_2 به‌عنوان دومین بیت است و روی این دو گروه آزمایش انجام می‌دهیم و این روند را برای تمام زوج‌های دیگر انجام می‌دهیم به‌طوری‌که دارای دقیقاً یک فرد بیمار باشیم. حال فرض می‌کنیم $(r + 1)$ -امین مرحله، اولین مرحله‌ای باشد که دارای دقیقاً دو نتیجه مثبت باشیم. فرض می‌کنیم گروه‌های $G_1^{j_1}, G_1^{j_2}, \dots, G_r^{j_r}$ در r مرحله اول دارای نتیجه مثبت هستند (در هر مرحله یک آزمون مثبت) و در مرحله $(r + 1)$ -ام هر دو گروه G_{r+1}^1 و G_{r+1}^0 مثبت‌اند. در این صورت دو گروه زیر را در نظر می‌گیریم:

گروه اول شامل تمامی افرادیست که $r + 1$ بیت اول آن‌ها ۰، $j_1, j_2, \dots, j_r$ است.

گروه دوم شامل تمامی افرادیست که $r + 1$ بیت اول آن‌ها ۱، $j_1, j_2, \dots, j_r$ است.

به وضوح هر گروه دارای 2^{n-r-1} نفر است که دقیقاً یک فرد بیمار دارند لذا بنا بر نتیجه (۴.۲.۳) با $n - r - 1$

آزمایش می‌توان فرد بیمار را شناسایی کرد بنابراین در این مرحله به $2(n - r - 1)$ آزمون مورد نیاز است و از آنجا

که در $r + 1$ مرحله اول $2(r + 1)$ آزمایش صورت می‌گیرد بنابراین با $2n = 2(r + 1) + 2(n - r - 1)$

آزمایش می‌توان حداکثر دو بیمار را از میان 2^n نفر شناسایی کرد.

فصل ۴

الگوریتم‌های آزمون گروهی ایستا و مدل آستانه

۱.۴ مقدمه

در این فصل ابتدا الگوریتم مدل معمولی آزمون گروهی معرفی می‌شود و همچنین الگوریتمی که با فرض وجود خطا می‌تواند تمام افراد بیمار را شناسایی کند، ارائه می‌دهیم. الگوریتم‌های مدل پوچگر و مدل پوچگر با خطا در بخش سوم این فصل مورد بررسی قرار می‌گیرند. در بخش چهارم الگوریتم‌های مدل هم‌تافت معرفی می‌شوند. در انتها آزمون گروهی آستانه^۱ را معرفی می‌کنیم. مدل آستانه تعمیمی از مدل معمولی آزمون گروهی است که اولین بار توسط دِمَشک^۲ مطرح شد [۷]. در این مدل نتیجه آزمون گروهی روی زیرمجموعه‌ای از کل افراد، مثبت است اگر شامل حداقل u بیمار باشد و منفی است اگر حداکثر l ($l < u$) بیمار در زیرمجموعه وجود داشته باشد.

۲.۴ الگوریتم‌های مدل معمولی

همانطور که در فصل اول اشاره شد، آزمون گروهی ایستا معمولاً با یک $(1, 0)$ - ماتریس معرفی می‌شود. به‌طوری‌که گروه‌ها، سطرهای ماتریس و افراد، ستون‌های ماتریس را تشکیل می‌دهند. یک گروه با خروجی مثبت (منفی) یک گروه مثبت (منفی) نامیده می‌شود. فرض می‌کنیم C^+ به عنوان ستون مثبت، متناظر با یک فرد بیمار باشد و همچنین C^- به عنوان ستون منفی، متناظر با یک فرد سالم باشد. با فرض اینکه بردار نتیجه R است، تعریف می‌کنیم:

^۱threshold group testing

^۲Damaschke

$$t_{\circ}^R(C) \equiv |C \setminus R|.$$

که $t_{\circ}^R(C)$ تعداد گروه‌هایی با خروجی منفی که فرد متناظر با ستون C در آن‌ها ظاهر می‌شود را نشان می‌دهد. به‌وضوح در آزمون گروهی ایستا برای هر ستون منفی C^- ، $t_{\circ}^R(C^-) \geq 1$ است و برای هر ستون مثبت C^+ ، $t_{\circ}^R(C^+) = 0$ است. همچنین تعریف می‌کنیم:

$$t_{\cap}^R(C) \equiv |C \cap R|.$$

که $t_{\cap}^R(C)$ تعداد گروه‌هایی با خروجی مثبت که فرد متناظر با ستون C در آن‌ها ظاهر می‌شود را نشان می‌دهد. الگوریتم زیر یک ابزار برای کد گشایی معمولی است.

ITEM SELECTION(N, R, D, e)

```

۰   for each item  $C \in N$ 
۱       compute  $t_{\circ}^R(C)$ 
۲       if  $t_{\circ}^R(C) \leq e$ 
۳           then  $D \leftarrow D \cup C$ 

۴   return  $D$ 
```

در الگوریتم فوق N ستون‌های متناظر با کل افراد شرکت‌کننده در آزمون گروهی را نشان می‌دهد و e حداکثر تعداد خطاهای اتفاق افتاده در بردار خروجی است. الگوریتم برای هر فرد مقدار $t_{\circ}^R(C)$ را بدست می‌آورد و مجموعه D برگردانده می‌شود. مجموعه D شامل همه افرادی است که، تحت خروجی مجموعه R ، در حداکثر e گروه مثبت ظاهر می‌شوند. در این فصل همه الگوریتم‌ها دارای الگوریتم فوق به‌عنوان زیر الگوریتم هستند، اما پارامترهای آن‌ها بسته به نیاز هر مدل متفاوت است. با فرض اینکه حداکثر d ستون مثبت در مجموعه N وجود دارد، الگوریتم d -پایه^۳ زیر را داریم.

^۳d-basic algorithm

$d - BASIC \ ALGORITHM$

- use a $d - disjunct \ matrix$
- ۱ $R \leftarrow the \ outcome - set$
- ۲ $D \leftarrow \emptyset$
- ۳ $ITEM \ SELECTION(N, R, D, \circ)$

قضیه ۱.۲.۴. [۶] الگوریتم d -پایه حداکثر d فرد بیمار را شناسایی می‌کند.

اثبات. واضح است $t_{\circ}^R(C^+) = \circ$ فرض می‌کنیم حداکثر d ستون مثبت در مجموعه N وجود دارد. ستون منفی C^- و اجتماع d ستون مثبت را در نظر می‌گیریم. بنابر خاصیت d -مجزایی حداقل یک سطر وجود دارد که ستون C^- در آن سطر دارای درایه‌ی یک ولی هر یک از این d ستون در این سطر دارای درایه‌ی صفر هستند. بنابراین خروجی این سطر منفی است پس $t_{\circ}^R(C^-) \geq 1$ لذا در شرط $t_{\circ}^R(C) = \circ$ فقط ستون‌های مثبت صدق می‌کنند. بنابراین D مجموعه‌ای با ستون‌های مثبت (افراد بیمار) است. $\square$

۱.۲.۴ الگوریتم مدل خطا آزاد

در این بخش الگوریتمی را ارائه می‌دهیم که با فرض وجود حداکثر e خطا در خروجی می‌تواند حداکثر d بیمار را شناسایی کند. همانطور که در فصل یک گفته شد یک ماتریس را (d, e) - مجزا می‌نامیم اگر برای هر $1 + d$ ستون $C_0, C_1, \dots, C_d$ داشته باشیم:

$$|C_0 \setminus \bigcup_{i=1}^d C_i| \geq e.$$

یعنی اینکه حداقل e سطر موجود است که در این سطرها C_0 دارای درایه یک است و هر ستون C_i ، $1 \leq i \leq d$ دارای درایه صفر است.

با فرض وجود حداکثر d بیمار و e خطا در خروجی الگوریتم (d, e) - **خطا آزاد** ^۴ زیر را داریم که به اختصار آن را با، الگوریتم $(d, e) - E$ نمایش می‌دهیم.

^۴ (d, e) - error-tolerant algorithm

$$(d, e) - E \text{ ALGORITHM}$$

- ۰ use a $(d; 2e + 1) - \text{disjunct matrix}$
- ۱ $R \leftarrow \text{the outcome - set}$
- ۲ $D \leftarrow \emptyset$
- ۳ $ITEM \text{ SELECTION}(N, R, D, e)$

قضیه ۲.۲.۴. [۶] یک الگوریتم (d, e) -خطا آزاد می‌تواند حداکثر d فرد بیمار را شناسایی کند.

اثبات. فرض می‌کنیم حداکثر e خطا در خروجی رخ دهد و همچنین حداکثر d ستون مثبت در مجموعه N وجود داشته باشد. ستون منفی C^- و اجتماع d ستون مثبت را در نظر می‌گیریم. برای هر ستون منفی C^- با استفاده از خاصیت $(d, 2e + 1)$ -مجزایی حداقل $2e + 1$ سطر وجود دارد که C^- در آن سطرها دارای درایه‌ی یک است ولی هر یک از d ستون مثبت دارای درایه‌ی صفر هستند. بنابراین نتیجه گروه‌های متناظر با این سطرها باید منفی باشند. بدترین حالت این است که e خطا در خروجی وجود داشته باشد که در این حالت نتیجه حداقل $e + 1$ تا از گروه‌ها منفی هستند، لذا $t_{\bullet}^R(C^-) \geq e + 1$. از طرفی با در نظر گرفتن حداکثر e خطا در خروجی برای هر ستون مثبت C^+ داریم $t_{\bullet}^R(C^+) \leq e$. بنابراین با استفاده از تابع $t_{\bullet}^R(C)$ می‌توانیم افراد بیمار و سالم را شناسایی کنیم. $\square$

۳.۴ مدل خطا آزاد با پوچگر (EI)

فرض می‌کنیم I مجموعه‌ای از پوچگرها باشد، که وجود حداقل یک پوچگر در یک گروه بدون در نظر گرفتن افراد بیمار نتیجه گروه را منفی می‌کند. با فرض وجود حداکثر h پوچگر در خروجی الگوریتم (d, h, e) - خطا آزاد با پوچگر^۵ زیر را داریم. یک الگوریتم (d, h, e) - خطا آزاد با پوچگر را به اختصار با، الگوریتم $(d, h, e) - EI$ نمایش می‌دهیم.

$(d, h, e) - EI$ ALGORITHM

```

۰ use a  $(d + h; \forall e + 1) -$  disjunct matrix
۱  $R \leftarrow$  the outcome - set
۲  $D \leftarrow \emptyset$ 
۳  $O \leftarrow \emptyset$ 
۴ for every item  $C \in N$ 
۵ compute  $t_1^R(C)$ 
۶ if  $t_1^R(C) \leq e$ 
۷ then  $O \leftarrow O \cup C$ 
۸ for all  $h -$  subsets  $S \subseteq O$ 
۹  $R \leftarrow R \cup (\bigcup_{C \in S} C)$ 
۱۰  $ITEM\ SELECTION(N \setminus O, R, D, e)$ 

```

دیاچکوف^۶ [۴۷] ابتدا یک الگوریتم ایستا را برای مدل پوچگر بدون خطا ارائه نمود و هدف اصلی را پیدا کردن خروجی‌های مثبت خنثی شده توسط پوچگرها قرار داد. در این الگوریتم تمام $\binom{n}{h'}$ زیرمجموعه از N را که $h \leq h'$ جستجو می‌کنیم. وانگ و لایو [۲۳] الگوریتم کدگشای بهتری را با در نظر گرفتن خطا معرفی کردند. در این الگوریتم تعداد جستجوها به $\binom{|O|}{h}$ کاهش پیدا می‌کند که O مجموعه همه پوچگرها و بدون فرد بیمار است. در این الگوریتم ابتدا $t_1^R(C)$ و $t_h^R(C)$ را برای هر فرد محاسبه می‌کنیم و سپس مجموعه N را به چهار زیرمجموعه تقسیم‌بندی می‌کنیم به‌طوری‌که پوچگرها از افراد بیمار جدا هستند. حال الگوریتم ساده‌ای را ارائه می‌دهیم که در

^۵ (d, e) - error-tolerant inhibitor algorithm

^۶ D'yachkov

این الگوریتم ابتدا مجموعه‌ی همه پوچگرها، O ، را جمع‌آوری می‌کنیم. ستون C را به‌عنوان ستون مثبت می‌شناسیم اگر $S \subseteq O$ وجود داشته باشد به‌طوری‌که $t_0^R(C) \leq e$ ، تحت بردار خروجی تنظیم شده توسط S .

قضیه ۱.۳.۴. [۶] الگوریتم $(d, h, e) - EI$ می‌تواند همه افراد بیمار را تحت مدل $(d, h, e) - EI$ شناسایی کند.

اثبات. نشان می‌دهیم الگوریتم $(d, h, e) - E$ برای مدل $(d, h, e) - E$ کار می‌کند. ابتدا نشان می‌دهیم O شامل همه پوچگرها است ولی شامل هیچ فرد بیماری نیست. بنابر خاصیت $(d + h, 2e + 1) -$ مجزایی هر شخصی که در حداکثر e گروه مثبت وجود داشته باشد بیمار نیست. فرض می‌کنیم ستون مثبت C^+ در حداکثر e گروه وجود داشته باشد. $d + h$ ستون $C^+ \setminus (P \cup I \cup C^-)$ را در نظر می‌گیریم که C^- یک ستون دلخواه سالم است. بنابر خاصیت $(d + h, 2e + 1) -$ مجزایی حداقل $2e + 1$ سطر وجود دارد که هر یک از $d + h$ ستون فوق در آن سطرها دارای درایه‌ی صفر هستند ولی ستون C^+ در آن سطرها دارای درایه‌ی یک است. از آنجایی که ستون C^+ مثبت است لذا خروجی گروه‌های بدون پوچگر متناظر با این سطرها مثبت هستند که این تناقض با حداکثر e گروه مثبت ظاهر شدن C^+ است. این موضوع حتی در بدترین حالت، با وجود حداکثر e خطا در خروجی برقرار است. بنابراین O شامل همه پوچگرها و بدون بیمار است.

ستون سالم $C^- \in N \setminus O$ و مجموعه P با حداکثر d ستون بیمار را در نظر می‌گیریم. برای هر زیرمجموعه h -عضوی S از O بنابر خاصیت $(d + h, 2e + 1) -$ مجزایی حداقل $2e + 1$ سطر وجود دارد به‌طوری‌که هر یک از $d + h$ ستون $K = P \cup S$ در این سطرها دارای درایه‌ی صفر هستند ولی C^- در این سطرها دارای درایه‌ی یک است. به‌وضوح باید خروجی گروه‌های متناظر با این سطرها با صرفنظر از خطا منفی باشد. حال با در نظر گرفتن حداکثر e خطا در خروجی داریم $1 + e = (2e + 1) - e = t_0^R(C^-)$. حال بنا بر الگوریتم $ITEM \ SELECTION(N \setminus O, R, D, e)$ هر C^- در شرط $t_0^R(C^-) \leq e$ صدق نمی‌کند بنابراین مجموعه D دارای هیچ فرد سالمی نیست.

زیرمجموعه h - عضوی S از O را در نظر می‌گیریم. هر فرد بیمار C^+ از $N \setminus O$ ، فقط در گروه‌هایی با خروجی جدید مجموعه $R \cup (\bigcup_{C \in S} C)$ ظاهر می‌شود. در بدترین حالت که e خطا در خروجی وجود دارد C^+ در حداکثر e گروه با نتیجه منفی ظاهر می‌شود. بنابراین هر فرد بیمار در مجموعه D ظاهر می‌شود. با توجه به مباحث فوق می‌توان نتیجه گرفت که خروجی الگوریتم $(d, h, e) - EI$ مجموعه‌ای با تمام افراد بیمار است. $\square$

۱.۳.۴ مدل معمولی خطا آزاد با پوچگر (GEI)

مدل بحث شده در بخش قبل ساده‌ترین نوع از پوچگرها است که وجود تنها یک پوچگر در یک گروه صرفنظر از وجود افراد بیمار در آن نتیجه منفی می‌دهد. [۴۶] این مدل را می‌توان به مدل k - پوچگر تعمیم داد به‌طوری که وجود k پوچگر در یک گروه نتیجه منفی می‌دهد. همچنین می‌توانیم فرضیات پیچیده‌تری را داشته باشیم، مثلاً می‌توان فرض کرد هر مجموعه با k پوچگر مجموعه‌ای با g فرد بیمار را خنثی می‌کند.

یک نتیجه که توسط چانگ^۷ [۴۲] مطرح شد این است که یک ماتریس $(d + h, 2e + 1)$ - مجزا تحت مدل $(d, h, e) - GEI$ به‌خوبی مدل $(d, h, e) - EI$ تمام افراد بیمار را شناسایی می‌کند. هدف این مدل این است که همه‌ی نتایج مثبت خنثی شده توسط پوچگرها را نمایش می‌دهد، اما در این مدل جدا کردن پوچگرها از افراد بیمار به‌صورت پیشرفته انجام نمی‌گیرد. بنابراین بجای جستجو کردن تمام زیرمجموعه‌های h - عضوی O تمام زیرمجموعه‌های h - عضوی N را جستجو می‌کنیم. الگوریتم معمولی (d, h, e) - خطا آزاد با پوچگر زیر را داریم که به‌اختصار آن را با، الگوریتم $(d, h, e) - GEI$ نمایش می‌دهیم.

$(d, h, e) - GEI$ ALGORITHM

- ۰ use a $(d + h; 2e + 1) - disjunct$ matrix
- ۱ $R \leftarrow$ the outcome - set
- ۲ $D \leftarrow \emptyset$
- ۳ for all h - subsets $S \subset N$
- ۴ $R \leftarrow R \cup (\bigcup_{C \in S} C)$
- ۵ $ITEM\ SELECTION(N \setminus S, R, D, e)$

^۷Chang

قضیه ۲.۳.۴. [۶] الگوریتم $GEI - (d, h, e)$ می‌تواند تمام افراد بیمار را تحت مدل $GE - (d, h, e)$ شناسایی کند.

اثبات. مشابه اثبات قضیه (۱.۳.۴) است. $\square$

۴.۴ الگوریتم‌های کدگشا برای مدل هم‌تافت خطا آزاد (EC)

در مدل EC مسأله را روی مدل هم‌تافت با درنظر گرفتن حداکثر e خطا در خروجی درنظر می‌گیریم. برای زیرمجموعه‌ی X از افراد تعریف می‌کنیم:

$$t_1^R(X) \equiv |(\cap X) \cap R| \quad t_\circ^R(X) \equiv |\cap X \setminus R|,$$

که تعداد گروه‌های مثبت و یا منفی را که هر یک از اعضای X در آن‌ها ظاهر می‌شوند را نمایش می‌دهند. $\cap X$ مجموعه گروه‌هایی را نشان می‌دهد که هر یک از اعضای X در آن‌ها وجود دارند.

وی^۸ و استینسون [۵۷] حالت خطا آزاد ماتریس $(d, r]$ - مجزا را ارائه دادند.

یک ماتریس دودویی $(d, r; z]$ - مجزا نامیده می‌شود اگر برای هر $d + r$ ستون $C_1, C_2, \dots, C_{d+r}$ داشته باشیم:

$$|\cap_{i=1}^r C_i \setminus \cup_{i=r+1}^{d+r} C_i| \geq z.$$

COMPLEX SELECTION(H, R, D, e)

- ۰ for each item $X \in H$
- ۱ compute $t_\circ^R(X)$
- ۲ if $t_\circ^R(X) \leq e$
- ۳ then $D \leftarrow D \cup X$
- ۴ return D

^۸Wei

با فرض اینکه پارامتر r حداکثر اندازه هر هم‌تافت را نشان می‌دهد الگوریتم $(d, r; e)$ - هم‌تافت خطا آزاد^۹ زیر را در نظر می‌گیریم که آن را به اختصار با، الگوریتم $(d, r; e) - EC$ نمایش می‌دهیم.

$(d, r; e) - EC \text{ ALGORITHM}$

- ۰ use a $(d, r; 2e + 1) - \text{disjunct matrix}$
- ۱ $R \leftarrow \text{the outcome set}$
- ۲ $D \leftarrow \emptyset$
- ۳ $\text{COMPLEX_SELECTION}(H, R, D, e)$

قضیه ۱.۴.۴. [۶] الگوریتم $(d, r; e) - EC$ می‌تواند هم‌تافت‌های مثبت را تحت مدل $(d, r; e) - EC$ شناسایی کند.

اثبات. بدترین حالت این است که e خطا در بردار خروجی وجود داشته باشد در این صورت $t_{\bullet}^R(X^+) \leq e$. هم‌تافت سالم X^- را در نظر می‌گیریم. مجموعه r - عضوی K شامل X^- و مجموعه d - عضوی T را چنان در نظر می‌گیریم که $T \cap K = \emptyset$ و با هر هم‌تافت مثبت دارای حداقل یک اشتراک باشد. با توجه به خاصیت $(d, r; 2e + 1)$ - مجزایی حداقل $2e + 1$ سطر وجود دارد که شامل K هستند اما شامل T نیستند. به وضوح هیچ یک از این $2e + 1$ گروه شامل هم‌تافت مثبت نیست، لذا نتیجه هر یک از این گروه‌ها منفی است. حتی در بدترین حالت که e خطا در خروجی وجود دارد X^- در حداقل $e + 1$ گروه منفی دیده می‌شود، یعنی $t_{\bullet}^R(X^+) \geq e + 1$.
 لذا با استفاده از این الگوریتم می‌توانیم هم‌تافت‌های مثبت را از هم‌تافت‌های منفی جدا کنیم. $\square$

۱.۴.۴ الگوریتم مدل هم‌تافت خطا آزاد با پوچگر (EIC)

در این بخش مدل هم‌تافت را با وجود خطا و وجود پوچگر معرفی می‌کنیم، با این فرض که در بین هم‌تافت‌هایی که زیرمجموعه افراد هستند حداکثر d هم‌تافت بیمار، حداکثر r - عضوی وجود دارد و همچنین فرض می‌کنیم حداکثر h پوچگر داریم و حداکثر e خطا در خروجی وجود دارد. ساده‌ترین مدل پوچگر را در نظر می‌گیریم. در زیر یک

^۹ $(d, r; e)$ - error-tolerant complex algorithm

الگوریتم رمزگشا که شبیه به الگوریتم $EI - (d, h, e)$ را ارائه می‌دهیم. با این تفاوت که افراد را با هم‌تافت‌ها جایگزین می‌کنیم.

الگوریتم (d, h, r, e) - **هم‌تافت خطا آزاد با پوچگر**^{۱۰} زیر را در نظر می‌گیریم که آن را به اختصار با، الگوریتم $EIC - (d, h, r, e)$ نمایش می‌دهیم.

$(d, h, r, e) - EIC \text{ ALGORITHM}$

- ۰ use a $(d + h, r; 2e + 1) - \text{disjunct matrix}$
- ۱ $R \leftarrow \text{the outcome} - \text{set}$
- ۲ $D \leftarrow \emptyset$
- ۳ $O \leftarrow \emptyset$
- ۴ for every complex $X \in H$
- ۵ compute $t_1^R(X)$
- ۶ if $t_1^R(X) \leq e$
- ۷ then $O \leftarrow O \cup X$
- ۸ for all $h - \text{subsets } S \subseteq O$
- ۹ $R \leftarrow R \cup (\bigcup_{S \in h} (\cap X))$
- ۱۰ $COMPLEX \text{ SELECTION}(H \setminus O, R, D, O, e)$

قضیه ۲.۴.۴. [۶] الگوریتم $EIC - (d, h, r, e)$ همه هم‌تافت‌های مثبت را تحت مدل $EIC - (d, h, r, e)$ شناسایی می‌کند.

اثبات. اثبات شبیه به قضیه (۱.۳.۴) است با این تفاوت که افراد را با هم‌تافت‌ها جایگزین می‌کنیم. ابتدا نشان می‌دهیم مجموعه O شامل تمام پوچگرها است به‌طوری‌که شامل هیچ هم‌تافت بیمار نیست. فرض می‌کنیم هم‌تافت بیمار X در حداکثر e گروه با نتیجه مثبت دیده شود. مجموعه r - عضوی R شامل هم‌تافت X را در نظر می‌گیریم. مجموعه‌ی $(d + h)$ - عضوی T شامل حداکثر h پوچگر را که $T \cap R = \emptyset$ را در نظر می‌گیریم. لذا بنابر خاصیت $(d + h, r; 2e + 1)$ - مجزایی حداقل $2e + 1$ سطر وجود دارد که شامل R هستند ولی شامل T نیستند. چون X هم‌تافت بیمار است پس باید خروجی گروه‌های متناظر با این $2e + 1$ سطر مثبت باشد. حتی با فرض وجود e

^{۱۰} $(d, r; e)$ - error-tolerant inhibitor complex algorithm

خطا در خروجی هم‌تافت X در حداقل $e + 1$ گروه با خروجی مثبت ظاهر می‌شود که این تناقض با این است که هم‌تافت X در حداکثر e گروه با خروجی مثبت دیده می‌شود.

هم‌تافت سالم $X^- \in H \setminus O$ را در نظر می‌گیریم. برای هر زیرمجموعه h - عضوی S از O بنا به خاصیت $(d + h, r; 2e + 1)$ - مجزایی حداقل $2e + 1$ سطر وجود دارد که در این سطرها اجتماع مجموعه $(d + h)$ - عضوی T ، که شامل مجموعه S و شامل اعضای یک مجموعه d - عضوی که با هر هم‌تافت مثبت دارای حداقل یک اشتراک است، دارای درایه‌ی صفر هستند ولی مجموعه r - عضوی R شامل X^- دارای درایه‌ی یک هستند. به‌وضوح خروجی گروه‌های متناظر با این $2e + 1$ سطر منفی هستند. لذا $t_*^R(X^-) = |\bigcap X^- \setminus R| \geq 2e + 1$. همچنین در بدترین حالت با در نظر گرفتن e خطا در خروجی داریم $t_*^R(X^-) \geq e + 1$ لذا مجموعه D هم‌تافت سالم ندارد.

زیرمجموعه h - عضوی S از O شامل حداکثر h پوچگر را در نظر می‌گیریم. هر هم‌تافت بیمار $X^+ \in H \setminus O$ در گروه‌هایی با خروجی مجموعه $(\bigcup_{X^+ \in S} (\cap X)) \cup R$ ظاهر می‌شود. در بدترین حالت که e خطا در خروجی وجود دارد این هم‌تافت در حداکثر e گروه منفی ظاهر می‌شود. یعنی $t_*^R(X^+) \leq e$. بنابراین هر هم‌تافت بیمار در مجموعه D است. $\square$

۵.۴ الگوریتم‌های ایستا برای آزمون گروهی آستانه

آزمون گروهی آستانه تعمیمی از آزمون گروهی معمولی است که این مدل جدید ابتدا توسط دمشک مطرح شد که مسأله به صورت زیر توصیف می‌شود.

اعداد غیرمنفی l و u که $l < u$ به ترتیب آستانه پایین و بالا نامیده می‌شوند. نتیجه آزمون گروهی برای زیرمجموعه S از افراد، مثبت است اگر زیرمجموعه S شامل حداقل u بیمار باشد و منفی است اگر S دارای حداکثر l بیمار باشد. اگر تعداد بیماران زیرمجموعه S بین l و u باشد نتیجه آزمون اختیاری است. به‌وضوح آزمون گروهی معمولی یک حالت خاص از آزمون گروهی آستانه با $l = 0$ و $u = 1$ است. در مدل آستانه کران پایین روی $|P|$ به عنوان فرض قبلی لازم است برای مثال اگر $|P| \leq u - 1$ در این صورت نتیجه آزمون روی تمام 2^n زیرمجموعه‌های ممکن از افراد منفی است لذا هیچ نتیجه‌ای را نمی‌توان گرفت. بنابراین در ادامه این فصل فرض می‌کنیم که $u \leq |P| \leq d$. الگوریتم‌هایی که برای آزمون گروهی معمولی برای شناسایی بیماران مطرح شد نمی‌تواند برای مدل آستانه استفاده شود. فرض می‌کنیم $g \equiv u - l - 1$ ^{۱۱} بین آستانه بالا و پایین باشد. برای فهمیدن و ساختن یک الگوریتم ایستای موثر برای مدل آستانه، با حالت خاص $g = 0$ کار می‌شود و همچنین آزمون گروهی آستانه را به حالت خطا آزاد گسترش می‌دهیم.

۱.۵.۴ حالت بدون اختلاف ($g = 0$) مدل آستانه

در این حالت یک ارتباط به ظاهر نامربوط بین حالت خاص مدل آستانه و مسأله جستجوی گراف ^{۱۲} پیدا می‌شود. به‌عنوان یک نتیجه، با استفاده از ماتریس‌های مجزا، یک الگوریتم ایستا برای مدل آستانه بدون اختلاف مطرح می‌شود. حال مسأله‌ی جستجوی گراف مطرح شده در این بخش را به‌صورت زیر ارائه می‌دهیم.

در مسأله‌ی جستجوی گراف هدف این است که گراف پنهان H تعریف شده روی مجموعه رأس‌های داده شده‌ی $V = \{1, 2, 3, \dots, n\}$ را با پرسیدن سوال‌هایی به‌صورت زیر بسازیم. برای زیرمجموعه K از V سوال این است

^{۱۱}gap

^{۱۲}graph search

که، آیا K شامل حداقل یک یال از H است؟

نتیجه یک گروه مثبت است اگر و تنها اگر K شامل حداقل یک یال از گراف H باشد. هدف این است تا همه یال‌های **گراف پنهان** H را با استفاده از مدل جستجوی فوق پیدا کنیم. فرض می‌کنیم $\|H\|$ تعداد یال‌ها در گراف H باشد که از بالا توسط ثابت d محدود می‌شود. در این بخش فرض می‌کنیم $|H|$ تعداد رأس‌های القا شده توسط یال‌های H است.

گراف‌هایی که در این بخش مطرح می‌شوند ابرگراف هستند. ابرگراف $H = (V, F)$ را در نظر می‌گیریم. مجموعه متناهی V رئوس ابرگراف نامیده می‌شود. $F = \{E_1, E_2, \dots, E_m\}$ خانواده‌ای از زیرمجموعه‌های V است به‌طوری که E_i ها یال‌های ابرگراف H نامیده می‌شوند.

به یک ابرگراف u -**ابرگراف**^{۱۳} گفته می‌شود اگر هر یال از ابرگراف شامل دقیقاً u رأس باشد. فرض می‌کنیم M زیرمجموعه‌ای از V باشد یک ابرگراف u -**کامل**^{۱۴} نسبت به M است اگر و تنها اگر هر زیرمجموعه u - عضوی از M یک یال از ابرگراف باشد.

قضیه ۱.۵.۴. [۵] فرض می‌کنیم در خروجی آزمون گروهی ایستا حداکثر e خطا وجود داشته باشد. هر u - ابرگراف $H = (V, F)$ که $|H| \leq d$ را می‌توانیم با استفاده از یک ماتریس $[1, u, 2e + 1, d - u + 1]$ - مجزا بسازیم.

اثبات. نتیجه هر گروه شامل حداقل یک یال از H بجز گروه‌هایی که در آن‌ها خطا رخ داده است، مثبت است. در بدترین حالت که نتیجه e تا از گروه‌ها اشتباه است برای هر زیرمجموعه u - عضوی $X^+ \in F$ داریم

$$t_e^R(X^+) \leq e$$

حال کافی است تا نشان دهیم برای هر زیرمجموعه u -عضوی $X \notin F$ ، $t_e^R(X) \geq e + 1$ ، فرض می‌کنیم M مجموعه رأس‌های القا شده توسط یال‌های F باشد. برای هر زیرمجموعه u - عضوی $X \notin F$ داریم:

^{۱۳} u -hypergraph

^{۱۴} u -complet

$$|X \cap M| \leq u - 1 \quad \text{یا} \quad |X \cap M| = u.$$

اگر $|X \cap M| \leq u - 1$ آن‌گاه زیرمجموعه $(d - u + 1)$ -عضوی Y متمایز از X را در نظر می‌گیریم به‌طوری‌که تعداد رأس‌هایی که در M وجود دارد ولی در Y وجود ندارد حداکثر $u - 1$ است.

اگر $|X \cap M| = u$ آن‌گاه فرض می‌کنیم Y زیرمجموعه $(d - u + 1)$ -عضوی متمایز از X باشد که شامل همه رأس‌های $M \setminus X$ است. در هر دو حالت فوق هیچ یال القا شده توسط رأس‌های خارج Y وجود ندارد. با استفاده از خاصیت $[d - u + 1, u, 2e + 1]$ -مجزایی حداقل $2e + 1$ سطر وجود دارد که شامل X است ولی شامل Y نیست. از آنجایی که گروه‌های متناظر با این سطرها شامل هیچ یالی از F نیستند پس حداقل $2e + 1$ گروه منفی شامل X وجود دارد. حتی در بدترین حالت که نتیجه e تا از گروه‌ها اشتباه باشد X در حداقل $e + 1$ گروه منفی ظاهر می‌شود یعنی $t_*^R(X) \geq e + 1$. از مباحث بالا می‌توان نتیجه گرفت که یال X متعلق به F است اگر و تنها اگر $t_*^R(X) \leq e$. $\square$

مجدداً به مدل آستانه بدون اختلاف برمی‌گردیم. به یاد داریم که هدف این است تا تمام بیماران (مجموعه P) مجموعه داده شده N را با آزمایش اینکه یک گروه شامل حداقل u بیمار است یا خیر، شناسایی کنیم. باقراردادن N به‌عنوان رئوس مجموعه V و خانواده تمام زیرمجموعه‌های u -عضوی از P به‌عنوان یال‌های مجموعه F ، مدل آستانه بدون اختلاف برای ساختن ابرگراف پنهان $H = (V, F)$ که نسبت به P ، u -کامل است، مدل مناسبی است. به‌وضوح ابرگراف $H = (V, F)$ که نسبت به P ، u -کامل باشد یک u -ابرگراف با $|H| \leq d$ است. زیرا اگر مجموعه M ، مجموعه رأس‌های القا شده توسط یال‌های H باشد آن‌گاه $P \subseteq M$. حال اگر x ای موجود باشد که $x \in M$ و $x \notin P$ آن‌گاه این تناقض با حداکثر d عضوی بودن M است.

قضیه ۲.۵.۴. [۵] برای مدل آستانه خطا آزاد و بدون اختلاف، یک ماتریس $[1, u, 2e + 1, d - u + 1]$ - مجزا می‌تواند برای شناسایی تمام افراد بیمار استفاده شود.

اثبات. ابرگراف $H = (V, F)$ که V برابر مجموعه افراد، N ، و خانواده F تمام زیرمجموعه‌های u - عضوی از P است، در نظر می‌گیریم. اگر برای هر زیرمجموعه u - عضوی X از N ، $t_0^R(X) \leq e$ ، آن‌گاه اعضای X متعلق به P هستند. این روند را ادامه داده تا اینکه مجموعه P بدست آید. $\square$

برای $u = 1$ و $e = 0$ نتایج معادل‌اند با نتایج یک ماتریس d - مجزا که می‌تواند برای شناسایی حداکثر d فرد بیمار روی مسأله آزمون گروهی معمولی بکار رود. الگوریتم زیر برای شناسایی مجموعه P به صورت زیر توصیف می‌شود.

الگوریتم ۱: (مدل آستانه خطا آزاد بدون اختلاف)

ابزار: یک ماتریس $[1, u, 2e + 1, d - u + 1]$ - مجزا.

کد گشایی: فرض می‌کنیم X متغیری باشد که نشان‌دهنده زیرمجموعه u - عضوی از N است و S مجموعه‌ای شامل اعضای X است، بقسمیکه $t_0^R(X) \leq e$. لذا

$$S = \{x : x \in X \text{ and } t_0^R(X) \leq e\}.$$

برای محاسبه مجموعه S کافی است تا برای هر زیرمجموعه u - عضوی X بررسی کنیم که $t_0(X) \leq e$.

قضیه ۳.۵.۴. [۵] برای مدل آستانه خطا آزاد و بدون اختلاف یک الگوریتم ایستا وجود دارد که بطور موفقیت‌آمیزی حداکثر d فرد بیمار را شناسایی می‌کند و نیاز به حداکثر

$$(2e + 1) \left(\frac{d+1}{u} \right)^u \left(\frac{d+1}{d-u+1} \right)^{d-u+1} [1 + (d+1)(1 + \ln(\frac{n}{d+1} + 1))],$$

آزمون دارد.

اثبات. بنا به قضیه (۲.۵.۴) با استفاده از یک ماتریس $t \times n$ ، $[1, u, 2e + 1, d - u + 1]$ - مجزا می‌توان با t آزمون مجموعه P را شناسایی کرد و بنابر نامساوی (۱.۱) حکم برقرار است. $\square$

۲.۵.۴ حالت عمومی مدل آستانه

در این بخش حالت عمومی مدل آستانه را در نظر می‌گیریم. در این حالت اختلاف $g \neq 0$ بین آستانه بالا و پایین وجود دارد. لذا نیاز به الگوریتم پیچیده‌تری داریم که در این حالت شناسایی مجموعه بیماران نسبت به مدل آستانه بدون اختلاف متفاوت است. دمشک نشان داد که پیدا کردن مجموعه P' با فرض $|P' \setminus P| \leq g$ و $|P \setminus P'| \leq g$ بهترین نتیجه ممکن است [۷]. برای حالت عمومی مدل آستانه سعی می‌کنیم تا مجموعه P' با خاصیت گفته شده در بالا را بجای شناسایی دقیق مجموعه P پیدا کنیم.

الگوریتم ۲: (مدل آستانه عمومی و خطا آزاد)

ابزار: استفاده از ماتریس $t \times n$, $[1, 2e + 1, u, d - l]$ - مجزا.

کد گشایی:

گام اول: ابرگراف $H_u = (V, F)$ را طوری می‌سازیم که $V = N$ و زیرمجموعه u -عضوی X از N یک یال در F است اگر و تنها اگر $t^R(X) \leq e$.

گام دوم: برای $i = 1, 2, \dots, m$ مجموعه رئوس p_i ها را که $|p_1| < |p_2| < \dots < |p_m|$ طوری قرار می‌دهیم که ابرگراف $H_u = (V, F)$ نسبت به p_i ها u -کامل باشد. p_1 اولیه را تمام u رأس یک یال اختیاری از گراف انتخاب می‌کنیم. برای پیدا کردن p_2 همه حالت‌های ممکن برای بدست آوردن مجموعه $(g + 1)$ -عضوی A از $V(H_u) \setminus p_1$ و مجموعه g -عضوی B از p_1 را بررسی می‌کنیم به طوری که H_u نسبت به $B \cup (p_1 \setminus A)$ ، u -کامل باشد. اگر چنین A و B وجود داشته باشد سپس قرار می‌دهیم $p_2 = (p_1 \setminus A) \cup B$. این روند را تا زمانی که p_m قابل گسترش نباشد و یا اینکه $|p_m| \geq d$ ادامه می‌دهیم. در این صورت خروجی مجموعه $p' = p_m$ است.

لم ۴.۵.۴ [۵] در الگوریتم ۲ بعد از گام اول برای هر زیرمجموعه u -عضوی X ، $X \in F$ دارای حداکثر g فرد خارج از P است. همچنین هر زیرمجموعه u -عضوی X^+ از P باید در F باشد.

اثبات. به وضوح هر زیرمجموعه u -عضوی X^+ از P در گروه‌هایی با نتیجه مثبت ظاهر می‌شود. حتی در بدترین

حالت که e خطا در خروجی وجود دارد، $t_e^R(X) \leq e$ است. بنابراین هر زیرمجموعه u -عضوی X^+ از P در F است.

حال نشان می‌دهیم هر زیرمجموعه u -عضوی X شامل حداقل $g + 1$ فرد خارج از P ، $t_e^R(X) > e$ است. هر زیرمجموعه u -عضوی X که دارای حداقل $g + 1$ فرد خارج از P باشد، دارای حداکثر l بیمار است. زیرا در این صورت زیرمجموعه X شامل حداکثر $l = u - (g + 1)$ بیمار از P است. حال زیرمجموعه $(d - l)$ -عضوی Y ، که مجزا از X است، را طوری انتخاب می‌کنیم که حداکثر l بیمار در Y نباشد. با استفاده از خاصیت $[d - l, u, 2e + 1]$ - مجزایی حداقل $2e + 1$ سطر وجود دارد که هر کدام شامل X هستند ولی شامل Y نیستند. بدون در نظر گرفتن نتایج اشتباه نتیجه گروه‌های متناظر با این سطرها باید منفی باشند، زیرا هر کدام از آنها شامل حداکثر l بیمار است. حتی در بدترین حالت که نتیجه e تا از گروه‌ها اشتباه است، X در حداقل $e + 1$ گروه منفی ظاهر می‌شود. بنابراین هر زیرمجموعه u -عضوی X با حداقل $g + 1$ فرد از P ، $t_e^R(X) > e$ است. $\square$

قضیه ۵.۵.۴. [۵] مجموعه P' در خروجی الگوریتم ۲ در شرط $|P' \setminus P| \leq g$ و $|P \setminus P'| \leq g$ صدق می‌کند.

اثبات. با استفاده از لم (۴.۵.۴) هر P_i که در گام دوم صدق می‌کند دارای حداکثر g فرد خارج از P است. بنابراین داریم $|P' \setminus P| \leq g$.

وقتی P' قابل گسترش نباشد یا $|P'| \geq d$ الگوریتم ۲ خاتمه می‌یابد. فرض می‌کنیم $|P'| \geq d$ با توجه به اینکه $|P' \setminus P| \leq g$ لذا P' دارای حداقل $d - g$ بیمار است. P دارای حداکثر d بیمار است لذا $P \setminus P'$ دارای حداکثر $d - (d - g) = g$ بیمار است و این یعنی $|P \setminus P'| \leq g$.

حال کافی است نشان دهیم که اگر P' قابل گسترش نباشد آن‌گاه $|P \setminus P'| \leq g$.

فرض می‌کنیم $|P \setminus P'| > g$. زیرمجموعه $g + 1$ -عضوی A از $P \setminus P'$ و مجموعه g -عضوی B که $P' \setminus P \subseteq B \subseteq P'$ را در نظر می‌گیریم. از آنجا که $|P' \setminus P| \leq g$ و P' اولیه دارای $u > g$ عضو است، لذا مجموعه B موجود است. بنابراین مجموعه $(P' \cup A) \setminus B$ زیرمجموعه P است. در نتیجه H_u نسبت به

□ $B \setminus (P' \cup A)$ ، u -کامل است که این تناقض دارد با اینکه P' قابل گسترش نیست.

نتیجه ۶.۵.۴. [۵] برای مدل آستانه عمومی خطا آزاد ماتریس $t \times n$ ، $[d-l, u, 2e+1]$ - مجزا می‌تواند برای

شناسایی مجموعه P' که $|P \setminus P'| \leq g$ و $|P' \setminus P| \leq g$ استفاده شود.

قضیه ۷.۵.۴. [۵] برای مدل آستانه عمومی خطا آزاد الگوریتم ایستا برای شناسایی مجموعه P' که $|P \setminus P'| \leq g$

و $|P' \setminus P| \leq g$ وجود دارد به‌طوری که حداکثر

$$(2e+1) \left(\frac{d+u-l}{u} \right)^u \left(\frac{d+u-l}{d-l} \right)^{d-l} [1 + (d+u-l)(1 + \ln(\frac{n}{d+u-l} + 1))],$$

تست استفاده می‌کند.

اثبات. بنا به قضیه (۶.۵.۴) با استفاده از یک ماتریس، $t \times n$ ، $[d-l, u, 2e+1]$ - مجزا می‌توان با t تست

مجموعه P' را شناسایی کرد و بنابر نامساوی (۱.۱) حکم برقرار است.

□

فصل ۵

ساختارهای ترکیباتی جدید و کاربردهای آن برای آزمون گروهی با پوچگرها

۱.۵ مقدمه

آزمون گروهی با پوچگرها یک نوع آزمون گروهی معمولی است که علاوه بر وجود افراد بیمار و سالم نوع دیگری از افراد به نام پوچگر در آن وجود دارد. در این مدل نتیجه آزمایش روی یک گروه مثبت است اگر و تنها اگر آن گروه شامل حداقل یک فرد بیمار و بدون پوچگر باشد. نتیجه آزمایش هر گروه شامل حداقل یک فرد پوچگر، صرفنظر از وجود افراد بیمار در آن گروه، منفی است.

در این فصل مسأله آزمون گروهی با پوچگرها را با دو حالت بررسی می‌کنیم. در حالت اول کران بالا از افراد بیمار را داریم و سپس یک کران پایین روی تعداد آزمون‌های مورد نیاز برای شناسایی افراد بیمار در آزمون گروهی ایستا را ارائه می‌دهیم. همچنین کران پایین بهینه روی تعداد آزمون‌های مورد نیاز برای شناسایی افراد بیمار در آزمون گروهی با پوچگرها را بدست می‌آوریم. در حالت دوم تعداد دقیق افراد بیمار را می‌دانیم.

دِبونز در آزمون گروهی با پوچگرها، زمانی که تعداد دقیق افراد بیمار را می‌دانیم، الگوریتم‌هایی با دو مرحله بطور کامل ایستا را به عنوان موثرترین الگوریتم‌های آزمون گروهی پویا معرفی کرد.

در این فصل ساختارهای ترکیباتی جدیدی را ارائه می‌دهیم. به عنوان حالت خاص کدهای ابرپوشش دهنده و انتخابگرها را ارائه می‌دهیم. در انتهای فصل در مورد خاصیت‌های مجزایی بعضی از ماتریس‌های شمول بحث

می‌کنیم. ماتریس شمول $\delta(n, d, k, \Gamma, r)$ در [۲۸، ۱۵] و ماتریس $\pi(q, n, d, k)$ در [۲۹] و ماتریس $\gamma(n, d, k)$ در [۳۸] مورد مطالعه قرار گرفت.

۲.۵ تعداد دقیق افراد بیمار را نمی‌دانیم.

در این بخش تعداد دقیق افراد بیمار را نداریم ولی می‌دانیم در مجموعه‌ی n عضوی N حداکثر d بیمار و حداکثر r پوچگر وجود دارد و مسأله‌ی آزمون گروهی به‌همراه پوچگر را، با الگوریتم‌هایی با یک مرحله بررسی می‌کنیم.

۱.۲.۵ الگوریتم‌هایی با یک مرحله

الگوریتم با یک مرحله یک الگوریتم کاملاً ایستا است که در آن همه گروه‌ها بطور همزمان مورد آزمایش قرار می‌گیرند.

مجموعه‌ی $C = \{C_1, C_2, \dots, C_n\}$ که از n بردار دودویی بطول t تشکیل شده است را یک کد دودویی به‌اندازه n و طول t می‌نامیم. هر C_j کدواژه نامیده می‌شود و برای هر $1 \leq i \leq t$ ، $C_j(i)$ بمعنی i -امین درایه‌ی C_j است. کد دودویی C می‌تواند با یک $(0, 1)$ -ماتریس، $t \times n$ ، $M = \|C_j(i)\|$ که $j = 1, 2, \dots, n$ و $i = 1, 2, \dots, t$ نمایش داده شود بطوری که ستون‌های ماتریس کدواژه‌های C هستند. لذا یک کد دودویی به‌صورت مجموعه‌ای از کدواژه‌ها یا $(0, 1)$ -ماتریس متناظر با آن نمایش داده می‌شود. در ادامه‌ی فصل مجموعه‌ی اعداد مثبت $\{1, 2, \dots, n\}$ را با $[n]$ نمایش می‌دهیم. برای هر $j \in [n]$ شخص j -ام را متناظر با ستون C_j از ماتریس M در نظر می‌گیریم. برای هر سطر i از ماتریس M زیرمجموعه‌ی $T_i = \{j \in [n] ; C_j(i) = 1\}$ را تعریف می‌کنیم. فرض می‌کنیم C یک کد دودویی به طول t باشد. $q > 1$ کد واژه‌ی (ستون‌های دودویی) $C_{l_1}, C_{l_2}, \dots, C_{l_q}$ را در نظر می‌گیریم. جمع دودویی این کدواژه‌ها را با $(C_{l_1} \cup C_{l_2} \cup \dots \cup C_{l_q})$ نمایش می‌دهیم. دو ستون دودویی C_j و C_h را در نظر می‌گیریم. ستون C_h به‌وسیله ستون C_j پوشیده می‌شود اگر برای هر i که $C_h(i) = 1$ آن‌گاه $C_j(i) = 1$.

برای هر دو مجموعه‌ی $Q \subseteq C$ و $J \subseteq C$ بطوری که $Q \cap J = \emptyset$ ، تعریف می کنیم:

$$R(Q, J) = (R(1), R(2), \dots, R(t)),$$

که برای هر $i = 1, 2, \dots, t$:

$$R(i) = \begin{cases} 1 & \text{و } \bigcup_{x \in J} x(i) = 0 \text{ و } \bigcup_{x \in Q} x(i) = 1 \\ 0 & \text{در غیر این صورت} \end{cases}.$$

برای $J = \emptyset$ داریم $R(Q, J) = \bigcup_{x \in Q} x(i)$. اگر $Q = \{C'\}$ که $C' \in C$ بردار $R(Q, J)$ را برای سادگی با $R(C', J)$ نمایش می دهیم که $J \subset (C \setminus C')$ است.

تعریف ۱.۲.۵. [۸] اعداد n, d و r را به طوری که $d + r \leq n$ در نظر می گیریم و همچنین $(0, 1)$ -ماتریس M با n ستون و t سطر را در نظر می گیریم. اگر برای زیرمجموعه‌های متمایز و حداکثر d -عضوی P_1 و P_2 و همچنین زیرمجموعه‌های متمایز و حداکثر r -عضوی I_1 و I_2 از ستون‌های ماتریس M به طوری که $P_1 \cap I_1 = \emptyset$ و $P_2 \cap I_2 = \emptyset$ داشته باشیم $R(P_1, I_1) \neq R(P_2, I_2)$ آن گاه به ماتریس M یک (d, r) -ساختار پوچگر^۱ می گوییم.

مینیم طول یک (d, r) -ساختار پوچگر به اندازه n را با $N_1(d, r, n)$ نمایش می دهیم.

ماتریس M یک الگوریتم یک مرحله‌ای است که بطور موفقیت آمیزی حداکثر d بیمار را با فرض وجود حداکثر r پوچگر شناسایی می کند، اگر و تنها اگر ماتریس M یک (d, r) -ساختار پوچگر باشد. زیرا اگر ماتریس M یک الگوریتم یک مرحله‌ای باشد که $R(P_1, I_1) = R(P_2, I_2)$ در این صورت بنابر قضیه (۲.۲.۳) باید $P_1 = P_2$ و همچنین اگر $P_1 \neq P_2$ و سایر شرایط برقرار باشد $R(P_1, I_1) \neq R(P_2, I_2)$.

بنابراین اگر ماتریس M یک (d, r) -ساختار پوچگر باشد آن گاه یک تناظر یک به یک بین بردارهای نتیجه و

^۱inhibitory (d,r)-design

حالت‌های ممکن از انتخاب ستون‌ها وجود دارد لذا ماتریس M یک الگوریتم یک مرحله‌ای است.

اگر P مجموعه ستون‌هایی از M متناظر با مجموعه‌ی افراد بیمار و I مجموعه ستون‌های متناظر با مجموعه افراد پوچگر باشد، آن‌گاه بردار نتیجه R برابر $R(P, I)$ است. زیرا اگر $R(P, I)(i) = 1$ باشد آن‌گاه $\bigcup_{x \in P} x(i) = 1$ و $\bigcup_{x \in I} x(i) = 0$ است و این یعنی سطر i -ام دارای حداقل یک فرد بیمار و بدون پوچگر است. لذا نتیجه گروه i -ام مثبت است بعبارت دیگر $R(i) = 1$ است. همچنین اگر $R(P, I)(i) = 0$ باشد آن‌گاه $\bigcup_{x \in P} x(i) = 1$ و $\bigcup_{x \in I} x(i) = 1$ است لذا سطر i -ام دارای شخص بیمار و پوچگر است یعنی $R(i) = 0$. برای حالت‌های بعد هم به همین ترتیب داریم.

یک (d, r) -ساختار پوچگر دلالت می‌کند براینکه برای هر مجموعه‌ی P' با حداکثر d ستون از M و برای هر مجموعه‌ی I' با حداکثر r ستون از M با فرض $P' \cap I' = \emptyset$ ، $R(P', I') = R(P, I)$ اگر و تنها اگر $P = P'$.

در ادامه یک کران پایین روی مینیم طول $N_1(d, r, n)$ از یک (d, r) -ساختار پوچگر را معرفی می‌کنیم.

قبل از معرفی این کران نیاز به تعریف کدهای (d, q) -ابرپوشش دهنده داریم. کدهای (d, q) -ابرپوشش دهنده دارای این خاصیت هستند که جمع هر d ستون به‌وسیله جمع هر q ستون دیگر پوشیده نمی‌شود. وقتی $d = 1$ این کدها با کدهای معمولی یا خانواده‌ی پوشش آزاد برابر است. مینیم طول یک (d, q) -ابرپوشش دهنده را با

$N(d, q, n)$ نشان می‌دهیم. که کران‌های پایین و بالای آن را به‌صورت زیر داریم [۴۸]:

$$N(d, q, n) = \begin{cases} \Omega\left(\frac{q^2}{d(\log q)} \log \frac{n}{q}\right) & q \geq 2d \\ \Omega(q \log \frac{n}{d}) & q < 2d \end{cases} \quad (1.5)$$

$$N(d, q, n) = O\left(\frac{q^2}{d} \log \frac{n}{q}\right). \quad (2.5)$$

قضیه ۲.۲.۵. [۸] برای هر عدد مثبت d, n و r با $n \geq d + r$ ، مینیم طول یک (d, r) -ساختار پوچگر به اندازه n به صورت زیر است:

$$N_1(d, r, n) \geq N(1, d + r - 1, n) = \Omega\left(\frac{(d+r)^2}{\log(d+r)} \log \frac{n}{d+r}\right).$$

اثبات. کافی است نشان دهیم که هر ماتریس (d, r) -ساختار پوچگر، M ، یک $(1, d + r - 1, n)$ -ابرپوشش دهنده است. فرض می کنیم ماتریس M یک (d, r) -ساختار پوچگر باشد ولی $(1, d + r - 1, n)$ -ابرپوشش دهنده نباشد. پس در ماتریس M ، $d + r$ ستون متمایز $C, C_1, C_2, \dots, C_{d+r-1}$ وجود دارد بطوری که ستون C توسط $C_1 \cup C_2 \cup \dots \cup C_{d+r-1}$ پوشیده می شود. مجموعه های $P = \{C_1, \dots, C_{d-1}\}$ و $I = \{C_d, \dots, C_{d+r-1}\}$ را در نظر می گیریم. اگر $C(i) = 0$ باشد سپس $R(C \cup P, I)(i) = R(P, I)(i)$ است. اگر $C(i) = 1$ باشد چون ستون C به وسیله ستون های $P \cup I$ پوشیده می شود، می توان نتیجه گرفت که $\bigcup_{C \in P} C(i) = 1$ یا $\bigcup_{C \in I} C(i) = 1$ و یا هر دو تساوی برقرار هستند. حال اگر $\bigcup_{C \in P} C(i) = 1$ در این صورت $R(C \cup P, I)(i) = R(P, I)(i)$ و لذا $R(C \cup P, I) = R(P, I)$ که این تساوی تناقض دارد با اینکه M یک (d, r) -ساختار پوچگر است. لذا با توجه به رابطه (۱.۵) حکم برقرار است. $\square$

قضیه ۳.۲.۵. [۸] فرض می کنیم مجموعه ای N شامل n فرد باشد، به طوری که شامل حداکثر $1 \leq d$ فرد بیمار و حداکثر $1 \leq r$ پوچگر باشد. هر الگوریتم یک مرحله ای که بطور موفقیت آمیزی افراد بیمار در N را شناسایی می کند حداقل $N(1, d + r - 1, n) = \Omega\left(\frac{(d+r)^2}{\log(d+r)} \log \frac{n}{d+r}\right)$ تست استفاده می کند.

اثبات. ماتریس M یک الگوریتم یک مرحله ای است، که بطور موفقیت آمیزی تمام افراد بیمار را با وجود حداکثر r پوچگر شناسایی می کند، اگر و تنها اگر M یک (d, r) -ساختار پوچگر باشد و از قضیه (۲.۲.۵) حکم برقرار است. $\square$

قضیه ۴.۲.۵. [۸] اعداد صحیح d, n و r با شرط $d + r \leq n$ را در نظر می گیریم. آن گاه یک (d, r) -ساختار پوچگر به اندازه n و طول $t = N(1, d + r, n) = O((d + r)^2 \log(\frac{n}{d+r}))$ وجود دارد.

اثبات. کافی است نشان دهیم یک ماتریس $(1, d + r, n)$ -ابرپوشش دهنده یک (d, r) -ساختار پوچگر است. زیرمجموعه های دلخواه، متمایز و حداکثر d -عضوی P_1 و P_2 و همچنین زیرمجموعه های دلخواه، متمایز و حداکثر r -عضوی I_1 و I_2 از ستون های ماتریس M را به طوری که $P_1 \cap I_1 = \emptyset$ و $P_2 \cap I_2 = \emptyset$ در نظر می گیریم. حال ستون $C \in P_2 \setminus P_1$ (یا $C \in P_1 \setminus P_2$) و مجموعه ستون های P_1 و I_2 با حداکثر $p + r$ عضو را در نظر می گیریم. بنا به خاصیت $(1, p + r, n)$ -ابرپوشش دهنده، لذا حداقل سطر i چنان وجود دارد که C در این سطر دارای درایه ای یک است و ستون های P_1 و I_2 در این سطر دارای درایه ای صفر هستند. لذا i -ای وجود دارد که $R(P_1, I_1)(i) \neq R(P_2, I_2)(i)$ و این یعنی $R(P_1, I_1) \neq R(P_2, I_2)$. $\square$

قضیه ۵.۲.۵. [۸] فرض می کنیم N مجموعه ای با n نفر و با حداکثر $1 \leq d$ بیمار و حداکثر $1 \leq r$ پوچگر باشد. یک الگوریتم یک مرحله ای وجود دارد که بطور موفقیت آمیزی می تواند تمام افراد بیمار در N را با حداکثر $t = N(1, d + r, n) = O((d + r)^2 \log(\frac{n}{d+r}))$ تست شناسایی کند.

اثبات. یک الگوریتم یک مرحله ای که تمام افراد بیمار را با وجود حداکثر r پوچگر شناسایی می کند موجود است اگر و تنها اگر (d, r) -ساختار پوچگر باشد. لذا بنا بر قضیه ی (۴.۲.۵) حکم برقرار است. $\square$

۳.۵ کدهای ابر پوشش دهندهی تعمیم یافته و انتخاب گرها

تعریف ۱.۳.۵. [۸] فرض می کنیم n و q دو عدد مثبت باشند. مجموعه‌ای از اعداد مثبت $P = \{p_1, p_2, \dots, p_s\}$ با فرض $1 \leq |P| = s$ و $q + \sum_{i=1}^s p_i \leq n$ را در نظر می گیریم. ماتریس M با n ستون و t سطر یک (P, q) -ابرپوشش دهنده است اگر برای هر $s+1$ زیرمجموعه‌های متمایز $A_1, A_2, \dots, A_s, B$ از ستون‌های M با $|B| = q$ و $i = 1, 2, \dots, s, |A_i| = p_i$ یک سطر از M موجود باشد به طوری که تمامی درایه‌های متناظر با ستون‌های B در این سطر برابر صفر باشند و برای هر $i \in \{1, 2, \dots, s\}$ درایه‌های متناظر با ستون‌های A_i شامل حداقل یک مقدار غیر صفر در این سطر باشند. n اندازه و t طول (P, q) -ابرپوشش دهنده است. مینیم طول یک (P, q) -ابرپوشش دهنده به اندازه n را با $N(P, q, n)$ نمایش می دهیم.

به توضوح برای $P = \{d\}$ تعریف (P, q) -ابرپوشش دهنده برابر تعریف (d, q) -ابرپوشش دهنده است و تعریف r -پوشش آزاد برابر تعریف $(\{1\}, r)$ -ابرپوشش دهنده است. سرانجام خانواده‌ی مجموعه‌هایی که فصل مشترک هر s تا از اعضای خانواده شامل اجتماع هر q عضو دیگر نباشد معادل است با (P, q) -ابرپوشش دهنده به طوری که P شامل s عدد برابر یک است.

در ادامه یک کران بالا برای مینیم طول $N(P, q, n)$ از یک (P, q) -ابرپوشش دهنده را مطرح می کنیم. برای این منظور ابتدا حالت تعمیم یافته‌ی (k, m, n) -انتخابگر را از دیونز ارائه می دهیم.

تعریف ۲.۳.۵. [۸] فرض می کنیم اعداد k, m, w, n و k که $1 \leq w < k \leq n$ و $1 \leq m \leq \binom{k}{w}$ را داریم. $(1, 0)$ -ماتریس M با t سطر و n ستون یک (k, m, w, n) -انتخابگر^۲ است اگر هر زیر ماتریس M با انتخاب k ستون اختیاری از ستون‌های M شامل حداقل m سطر با دقیقاً w درایه‌ی برابر یک باشد. عدد t را اندازه (k, m, w, n) -انتخابگر می نامیم. مینیم اندازه یک (k, m, w, n) -انتخابگر را با $t(k, m, w, n)$ نمایش می دهیم.

^۲ (k, m, w, n) -selector

می‌توان دید که یک (k, m, w, n) -انتخابگر با پارامترهای $k = q + \sum_{i=1}^s p_i$ و $m = \binom{k}{w} - \prod_{i=1}^s p_i + 1$ و $w = s$ یک $(\{p_1, p_2, \dots, p_s\}, q)$ -ابرپوشش دهنده است. لذا با بدست آوردن یک کران بالا روی اندازه انتخابگرهای تعمیم یافته بطول n می‌توان یک کران بالا روی مینیم طول $(p_1, p_2, \dots, p_s, q)$ -ابرپوشش دهنده به اندازه n بدست آورد.

قضیه ۳.۳.۵. [۸] برای هر عدد k, m, w, n به‌طوری که $k \leq n$ و $1 \leq w < k$ و $1 \leq m \leq \binom{w}{k}$ و $n \geq k^2 w$ یک (k, m, w, n) -انتخابگر به اندازه t وجود دارد با:

$$t < \frac{ek^{w+1}}{z} \ln \left\lceil \frac{n}{k} \right\rceil - \frac{ewkw}{z} \ln w + \frac{ek^w}{z} (w + m + k + z - 1),$$

که $ez = \binom{k}{w} - m + 1$ عدد نپر است.

اثبات. فرض می‌کنیم X یک مجموعه‌ی متناهی باشد و F یک خانواده از زیرمجموعه‌های X باشد. ابرگراف $H = (X, F)$ را در نظر می‌گیریم که X مجموعه‌ی رأس‌ها و F مجموعه‌ی ابريال‌ها است. زیرمجموعه‌ی $T \subseteq X$ را یک پوشش از H می‌نامیم اگر برای هر $E \in F$ ، $T \cap E \neq \emptyset$. مینیم اندازه‌ی یک پوشش از ابرگراف H را با $\tau(H)$ نمایش می‌دهیم. لواژ^۳ کران بالای $\tau(H)$ از پوشش H را به‌صورت زیر ارائه نمود [۲۷]:

$$\tau(H) < \frac{|X|}{\min_{E \in F} |E|} (1 + \ln \Delta), \quad (3.5)$$

ز که

$$\Delta = \max_{x \in X} |\{E : E \in F, x \in E\}|.$$

فرض می‌کنیم X مجموعه‌ی بردارهای دودویی $x = (x(1), x(2), \dots, x(n))$ به‌طول n با $b \geq w$ درایه‌ی برابر یک باشد و همچنین فرض می‌کنیم U مجموعه‌ی بردارهای دودویی به‌طول k با w درایه‌ی برابر یک باشد. برای

^۳Lovász

هر زیرمجموعه از اندیس‌های $s = \{i_1, i_2, \dots, i_k\}$ با $1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n$ و برای هر بردار دودویی $u = (u(1), u(2), \dots, u(k)) \in U$ مجموعه بردارهای دودویی زیر را تعریف می‌کنیم:

$$E_{u,s} = \{x = (x(1), x(2), \dots, x(n)) \in X : x(i_1) = u(1), \dots, x(i_k) = u(k)\}.$$

برای هر مجموعه‌ی $A \subseteq U$ به اندازه r که $r = 1, 2, \dots, \binom{k}{w}$ و هر مجموعه‌ی $s \subseteq \{1, 2, \dots, n\}$ با $|s| = k$ تعریف می‌کنیم:

$$F_r = \{E_{A,s} : A \subset U, |A| = r, s \subseteq \{1, 2, 3, \dots, n\}, |s| = k\},$$

که

$$E_{A,s} = \bigcup_{u \in A} E_{u,s}.$$

برای هر $r = 1, 2, \dots, \binom{k}{w}$ ابرگراف $H_r = (X, F_r)$ را در نظر می‌گیریم. فرض می‌کنیم $z = \binom{k}{w} - m + 1$ و نشان می‌دهیم که هر پوشش T از H_z یک (k, m, w, n) -انتخابگر است. فرض می‌کنیم T یک پوشش از H_z باشد به طوری که زیرماتریسی از T با ستون‌های $s = \{i_1, i_2, \dots, i_k\}$ وجود دارد که دارای حداکثر $m - 1$ سطر متمایز، با w درایه‌ی برابر یک است. فرض می‌کنیم $u_{j_1}, u_{j_2}, \dots, u_{j_q}$ که $q \leq m - 1$ چنین سطریابی باشند. بنابراین زیرمجموعه‌ی A از U به اندازه $|A| = z = \binom{k}{w} - m + 1$ وجود دارد که شامل هیچ یک از بردارهای $u_{j_1}, u_{j_2}, \dots, u_{j_q}$ نیست. یعنی اینکه T شامل هیچ راسی از یال‌های متناظر $E_{A,s}$ از H_z نیست و این تناقض دارد با فرض اینکه T یک پوشش برای H_z است. بنابراین T یک (k, m, w, n) -انتخابگر است. حال از نامساوی (۳.۵) استفاده می‌کنیم تا اینکه یک کران بالا روی t را مشخص کنیم. برای این منظور نیاز به محاسبه کمیت‌های زیر، برای ابرگراف H_z ، داریم:

$$\min\{|E| : E \in F_z\} \quad , \quad \Delta \quad \text{و} \quad |X|.$$

با توجه به اینکه مجموعه‌ی X شامل تمام بردارهای دودویی بطول n با b درایه برابر یک است، بنابراین $|X| = \binom{n}{b}$ است.

به منظور محاسبه $\min\{|E| : E \in F_z\}$ ، می‌دانیم که هر ابريال $E_{A,s} \in F_z$ اجتماع z تا مجموعه‌ی متمایز $E_{u,s}$ است و هر مجموعه‌ی $E_{u,s}$ شامل $\binom{n-k}{b-w}$ بردار است. بنابراین هر ابريال $E_{A,s}$ دارای اندازه $|E_{A,s}| = z \binom{n-k}{b-w}$ است.

حال باید Δ را محاسبه کنیم. به وضوح برای هر $x \in X$ ، $\binom{b}{w} \binom{n-b}{k-w}$ مجموعه‌های متمایز $E_{u,s}$ شامل x وجود دارند و برای هر $u \in U$ ، $\binom{z+m-2}{z-1}$ زیرمجموعه‌های متمایز A از U به اندازه z شامل u وجود دارند. در نتیجه هر $x \in X$ به $\Delta = \binom{b}{w} \binom{n-b}{k-w} \binom{z+m-2}{z-1}$ ابريال تعلق دارد.

بنابراین از رابطه (۳.۵) داریم:

$$t = \tau(H_z) < \frac{\binom{n}{b}}{z \binom{n-k}{b-w}} \left[1 + \ln \left(\binom{b}{w} \binom{n-b}{k-w} \binom{z+m-2}{z-1} \right) \right]. \quad (۴.۵)$$

برای سادگی فرض می‌کنیم n مضربی از k باشد و در نامساوی (۴.۵)، $b = \frac{n}{k}$ را قرار می‌دهیم.

ابتدا کران بالا روی

$$\frac{\binom{n}{\frac{n}{k}}}{\binom{n-k}{\frac{n}{k}-w}},$$

را پیدا می‌کنیم که برای $k = 2$ داریم:

$$\frac{\binom{n}{\frac{n}{k}}}{\binom{n-k}{\frac{n}{k}-w}} = \frac{\binom{n}{\frac{n}{2}}}{\binom{n-2}{\frac{n}{2}-1}} < 2e.$$

و برای $k \geq 3$ داریم:

$$\frac{\binom{n}{\frac{n}{k}}}{\binom{n-k}{\frac{n}{k}-w}} = k \left(\prod_{i=1}^{k-w} \frac{n-i}{n-\frac{n}{k}-i+1} \right) \left(\prod_{i=1}^{w-1} \frac{n-k+w-i}{\frac{n}{k}-i} \right)$$

که

$$\frac{\binom{n}{\frac{n}{k}}}{\binom{n-k}{\frac{n}{k}-w}} \leq k \left(\frac{n-k+w}{n-\frac{n}{k}-k+w+1} \right)^{k-w} \left(\frac{n-k+1}{\frac{n}{k}-w+1} \right)^{w-1}$$

$$\begin{aligned}
 &= k \left(\frac{n-k+w}{n-\frac{n}{k}-k+w+1} \right)^{k-1} \left(\frac{n-k+w}{n-\frac{n}{k}-k+w+1} \right)^{-(w-1)} \\
 &\quad \times \left(\frac{n-k+1}{\frac{n}{k}-w+1} \right)^{w-1} \\
 &= k \left(\frac{k(n-k+w)}{k(n-k+w)-(n-k)} \right)^{k-1} \\
 &\quad \times \left(\frac{(n-\frac{n}{k}-k+w+1)(n-k+1)}{(n-k+w)(\frac{n}{k}-w+1)} \right)^{w-1}. \tag{۵.۵}
 \end{aligned}$$

و از طرفی داریم:

$$\begin{aligned}
 \left(\frac{k(n-k+w)}{k(n-k+w)-(n-k)} \right)^{k-1} &= \left(1 + \frac{n-k}{k(n-k+w)-(n-k)} \right)^{k-1} \\
 &< \left(1 + \frac{1}{k-1} \right)^{k-1} < e.
 \end{aligned}$$

و همچنین داریم:

$$\begin{aligned}
 &\left(\frac{(n-\frac{n}{k}-k+w+1)(n-k+1)}{(n-k+w)(\frac{n}{k}-w+1)} \right)^{w-1} \\
 &= k^{w-1} \left(\frac{(n-\frac{n}{k}-k+w+1)(n-k+1)}{(n-k+w)(n-kw+k)} \right)^{w-1} \\
 &\leq k^{w-1} \left(\frac{n-\frac{n}{k}-k+w+1}{n-kw+k} \right)^{w-1}
 \end{aligned}$$

از آنجا که $n \geq k^2 w$

$$\begin{aligned}
 &\leq k^{w-1} \left(\frac{n-kw-k+w+1}{n-kw+k} \right)^{w-1} \\
 &< k^{w-1}
 \end{aligned}$$

بنابراین داریم:

$$\frac{\binom{n}{\frac{n}{k}}}{\binom{n-k}{\frac{n}{k}-w}} < ek^w, \tag{۶.۵}$$

با توجه به نامساوی $\binom{a}{b} \leq (\frac{ea}{b})^b$ داریم:

$$\begin{aligned}
 \Delta &= \binom{\frac{n}{k}}{w} \binom{n-\frac{n}{k}}{k-w} \binom{z+m-2}{z-1} \\
 &\leq \left(\frac{n}{kw}\right)^w \left(\frac{n-\frac{n}{k}}{k-w}\right)^{k-w} \left(\frac{z+m-2}{z-1}\right)^{z-1} e^{k+z-1} \\
 &= \left(\frac{n}{k}\right)^k \left(\frac{1}{w}\right)^w \left(\frac{k-1}{k-w}\right)^{k-w} \left(\frac{z+m-2}{z-1}\right)^{z-1} e^{k+z-1} \\
 &< \left(\frac{n}{k}\right)^k \left(\frac{1}{w}\right)^w \left(\frac{k}{k-w}\right)^{k-w} \left(\frac{z+m-2}{z-1}\right)^{z-1} e^{k+z-1} \\
 &= \left(\frac{n}{k}\right)^k \left(\frac{1}{w}\right)^w \left(1 + \frac{w}{k-w}\right)^{k-w} \left(1 + \frac{m-1}{z-1}\right)^{z-1} e^{k+z-1}.
 \end{aligned}$$

از آنجایی که $e^w \leq \left(1 + \frac{w}{k-w}\right)^{k-w}$ و $e^{m-1} \leq \left(1 + \frac{m-1}{z-1}\right)^{z-1}$ داریم:

$$\Delta \leq \left(\frac{n}{k}\right)^k \left(\frac{1}{w}\right)^w e^w e^{m-1} e^{k+z-1}. \quad (7.5)$$

از نامعادلات (۴.۵)، (۶.۵) و (۷.۵) داریم:

$$\tau = \tau(H_z) < \frac{ek^w}{z} \left(1 + \ln \left(\left(\frac{n}{k}\right)^k \left(\frac{1}{w}\right)^w e^{w+m+k+z-2} \right)\right).$$

اگر n مضربی از k نباشد آن گاه قرار می‌دهیم $n' = k \lceil \frac{n}{k} \rceil$ ، از مبحث بالا یک (k, m, w, n') -انتخابگر وجود دارد

که:

$$t < \frac{ek^w}{z} \left(1 + \ln \left(\lceil \frac{n}{k} \rceil^k \left(\frac{1}{w}\right)^w e^{w+m+k+z-2} \right)\right).$$

□

۴.۵ یک الگوریتم با دو گام برای مسأله آزمون گروهی با پوچگر

[۸] در این بخش یک الگوریتم با دو گام را برای مسأله آزمون گروهی با پوچگر معرفی می‌کنیم که فرض این

است که تعداد دقیقی از افراد بیمار را می‌دانیم. این دو گام به صورت زیر توصیف می‌شود:

گام اول: در این گام مجموعه‌ی حداکثر $(2r + 2p - 2)$ -عضوی شامل همه‌ی پوچگرها و افراد بیمار را

شناسایی می‌کنیم. فرض می‌کنیم M_1 یک $(\{p, r\}, r)$ -ابرپوشش دهنده و M_2 یک $(p, 2r + p - 1)$ -ابرپوشش

دهنده باشد. در این گام همه‌ی آزمون‌های مربوط به سطرهای M_1 و M_2 را بطور همزمان انجام می‌دهیم. نشان

می‌دهیم که نتایج آزمون‌های مربوط به سطرهای M_1 مجموعه‌ای شامل افراد پوچگر را مشخص می‌کند. متمم، بردار

نتیجه‌ی M_1 را با w نمایش می‌دهیم که i -امین درایه‌ی آن برابر یک است اگر و تنها اگر نتیجه آزمون مربوط به

i -امین سطر از M_1 منفی باشد و در غیر این صورت صفر است. مجموعه ستون‌هایی از M_1 را که با بردار w پوشیده

می‌شوند با X نمایش می‌دهیم. فرض می‌کنیم x ستونی از M_1 متناظر با شخصی در I باشد اگر i -امین درایه از

x برابر یک باشد این بدین معنی است که i -امین گروه شامل حداقل یک پوچگر است و در نتیجه پاسخ i -امین

تست منفی است و این یعنی $w(i) = 1$ است. در نتیجه در ماتریس M_1 برای هر ستون x ، متناظر با شخصی در I ،

$x \in X$ است. به وضوح X شامل بعضی از ستون‌های متناظر با افراد بیمار است. نشان می‌دهیم که ستون‌هایی از X

که متناظر با افراد بیمار نیستند حداکثر $2r - 1$ است، یعنی $|X \setminus P| \leq 2r - 1$. فرض می‌کنیم $|X \setminus P| \geq 2r$.

از آنجایی که $|I| \leq r$ لذا زیرمجموعه‌ی r -عضوی A از $X \setminus P$ وجود دارد به طوری که شامل ستون‌های متناظر

با افراد I نیست. از آنجایی که M_1 یک $(\{p, r\}, r)$ -ابرپوشش دهنده است سطر i -ام چنان وجود دارد که همه‌ی

ستون‌های متناظر با افراد I در این سطر دارای درایه‌ی برابر صفر هستند و حداقل یک ستون از P و یک ستون از

A در این سطر دارای درایه‌ی برابر یک هستند و این دلالت می‌کند بر اینکه پاسخ تست i -ام، شامل حداقل یک

فرد بیمار و بدون پوچگر، مثبت است. بنابراین i -امین درایه‌ی بردار w برابر صفر است لذا حداقل یک ستون از A

توسط بردار w پوشیده نمی‌شود. حال باید مجموعه‌ای به اندازه‌ی کمتر از $|P \setminus X| + p$ شامل همه‌ی افراد بیماری

که با ستون‌های X اشتراکی ندارد را مشخص کنیم. فرض می‌کنیم z بردار نتیجه‌ی تست‌های مربوط به سطرهای M_2 باشد و همچنین فرض می‌کنیم Y مجموعه ستون‌هایی به صورت زیر باشد:

$$Y = \{C : \text{ستونی از } M_2 \text{ است که } C \notin X \text{ و } R(C, X) \text{ با } z \text{ پوشیده می‌شود}\}.$$

اگر ستون دلخواه $C \notin X$ متناظر با یک فرد بیمار باشد آن‌گاه $R(C, I)$ توسط z پوشیده می‌شود. زیرا اگر $R(C, I)(i) = 1$ باشد یعنی اینکه ستون C ، متناظر با یک فرد بیمار، در سطر i -ام برابر یک است و در این سطر ستون‌های I برابر صفر هستند در این صورت پاسخ این تست مثبت است یعنی $R(i) = 1$ است.

نشان دادیم که $I \subseteq X$ ، بنابراین به ازای هر ستون $C \notin X$ ، $R(C, X)$ توسط $R(C, I)$ پوشیده می‌شود. زیرا اگر $R(C, X)(i) = 1$ باشد آن‌گاه ستون C در سطر i دارای درایه یک است و تمام ستون‌های X ، که I را هم شامل‌اند، در این سطر برابر صفر هستند لذا $R(C, I)(i) = 1$ است. در نتیجه اگر $C \notin X$ ستون متناظر با یک فرد بیمار باشد سپس بردار $R(C, X)$ توسط z پوشیده می‌شود لذا $P \setminus X \subseteq Y$.

ادعا می‌کنیم $|Y| < p + |P \setminus X|$ است. فرض می‌کنیم $|Y| \geq p + |P \setminus X|$ باشد. زیرمجموعه‌ی p -عضوی از L از $Y \setminus P$ را در نظر می‌گیریم، از آنجا که $|Y \cap P| \leq |P \setminus X|$ مجموعه‌ی L وجود دارد. از آنجا که M_2 یک $(p, p + 2r - 1)$ -ابرپوشش دهنده است و همچنین $|X \cup P| \leq p + 2r - 1$ ، لذا سطر i مانند i وجود دارد به طوری که حداقل یک ستون از L دارای i -امین درایه‌ی برابر یک است و همه‌ی ستون‌های $X \cup P$ در این سطر دارای درایه‌ی برابر صفر هستند. در نتیجه i -امین گروه دارای هیچ فرد بیماری نیست لذا $z(i) = R(P, I)(i) = 0$. به عبارت دیگر حداقل یک ستون $C' \in L \subset Y$ وجود دارد که در سطر i -ام دارای درایه یک است ولی در این سطر همه ستون‌های X برابر صفر هستند پس $R(C', X)(i) = 1$ که این تناقض دارد با اینکه هر $y \in Y$ توسط z پوشیده می‌شود.

گام دوم: در این گام به صورت آزمون انفرادی افراد متناظر به ستون‌های $X \cup Y$ را مورد آزمایش قرار می‌دهیم. آن‌گاه افراد بیمار را اعلام می‌نماییم.

الگوریتم فوق نتیجه زیر را به دنبال دارد.

قضیه ۱.۴.۵. [۸] فرض می‌کنیم که O مجموعه‌ای از n فرد باشد که شامل دقیقاً $1 \leq p$ فرد بیمار و حداکثر $1 \leq r$ پوچگر است. یک الگوریتم دو مرحله‌ای وجود دارد که بطور موفقیت آمیزی تمام افراد بیمار در O را به وسیله‌ی حداکثر $2p + 2r - 2$ تست $N(\{p, r\}, r, n) + N(p, p + 2r - 1, n)$ شناسایی می‌کند.

اثبات. با توجه به ساختار الگوریتم فوق، ماتریس M_1 نیاز به $N(\{p, r\}, r, n)$ تست دارد و ماتریس M_2 نیاز به $N(p, p + 2r - 1, n)$ تست دارد و همچنین مجموعه‌ی $X \cup Y$ نیاز به حداکثر $2p + 2r - 2$ تست دارد که حاصل جمع آن‌ها حداکثر تعداد تست‌هایی است که می‌توان با الگوریتم دو مرحله‌ای تمام افراد بیمار را شناسایی کرد. $\square$

۵.۵ ماتریس‌های شمول

مجموعه‌ی k -عضوی از $[n]$ را k -مجموعه ^۴ می‌نامیم و تمام k -مجموعه‌ها از $[n]$ را با $\binom{[n]}{k}$ نمایش می‌دهیم. فرض می‌کنیم Γ یک خانواده از k -مجموعه‌ها روی $[n]$ باشد به‌طوری‌که به ازای هر k -مجموعه K_1 و K_2 از Γ ، $|K_1 \setminus K_2| \geq r$ باشد. به عبارت دیگر بردارهای وقوع (بطول n) اعضای Γ یک کد دودویی با مینیمم فاصله همینگ $2r$ را تشکیل می‌دهند.

تعریف ۱.۵.۵. [۱۶] فرض می‌کنیم $1 \leq d \leq k \leq n$ و Γ یک خانواده از k -مجموعه‌ها روی $[n]$ با مینیمم فاصله همینگ $2r$ باشد. $(0, 1)$ -ماتریس $\delta(n, d, k, \Gamma, r)$ از مرتبه $\binom{[n]}{d} \times |\Gamma|$ را در نظر می‌گیریم به‌طوری‌که سطرهاى آن را اعضای $\binom{[n]}{d}$ و ستون‌های آن را اعضای Γ تشکیل می‌دهند.

برای $D \in \binom{[n]}{d}$ و $K \in \Gamma$ درایه‌ی (D, K) -ام از ماتریس $\delta(n, d, k, \Gamma, r)$ برابر یک است اگر و تنها اگر $D \subseteq K$ باشد.

^۴k-set

جدول ۱.۵: $\delta(4, 2, 3)$				
$\delta(4, 2, 3)$	$\{1, 2, 3\}$	$\{1, 2, 4\}$	$\{1, 3, 4\}$	$\{2, 3, 4\}$
$\{1, 2\}$	۱	۱	۰	۰
$\{1, 3\}$	۱	۰	۱	۰
$\{1, 4\}$	۰	۱	۱	۰
$\{2, 3\}$	۱	۰	۰	۱
$\{2, 4\}$	۰	۱	۰	۱
$\{3, 4\}$	۰	۰	۱	۱

اگر $\Gamma = \binom{[n]}{k}$ ، $(0, 1)$ -ماتریس $\delta(n, d, k, \Gamma, 1)$ را برای سادگی با $\delta(n, d, k)$ نمایش می‌دهیم. به‌وضوح وزن سطر و ستون یک ماتریس $\delta(n, d, k)$ به‌ترتیب $\binom{n-d}{k-d}$ و $\binom{k}{d}$ است. درجدول (۱.۵) یک $\delta(4, 2, 3)$ را داریم. برای $n \leq k$ ، مجموعه‌ی همه‌ی بردارهای q -تایی بطول n و وزن k را با $\binom{[n]}{k}[q]^k$ نمایش می‌دهیم. $\alpha \in \binom{[n]}{d}[q]^d$ و $\tau \in \binom{[n]}{k}[q]^k$ را درنظر می‌گیریم، $\alpha(i)$ و $\tau(i)$ را به‌ترتیب i -امین درایه از α و i -امین درایه از τ می‌نامیم. اگر برای $1 \leq i \leq n$ و $\alpha(i) \neq 0$ داشته‌باشیم $\alpha(i) = \tau(i)$ آن‌گاه $\alpha \prec \tau$ است. برای مثال فرض می‌کنیم $03010 \in \binom{[5]}{4}[3]^4$ و $03112 \in \binom{[5]}{4}[3]^4$ لذا $03112 \prec 03010$ است.

تعریف ۲.۵.۵. [۱۶] فرض می‌کنیم $1 \leq d \leq k \leq n$ و $q \geq 1$. $(0, 1)$ -ماتریس $\pi(q, n, d, k)$ از مرتبه $\binom{[n]}{d}q^d \times \binom{[n]}{k}q^k$ را درنظر می‌گیریم، به‌طوری‌که سطرهای آن را اعضای $\binom{[n]}{d}q^d$ و ستون‌های آن را اعضای $\binom{[n]}{k}q^k$ تشکیل می‌دهند.

برای $\alpha \in \binom{[n]}{d}q^d$ و $\tau \in \binom{[n]}{k}q^k$ درایه‌ی (α, τ) -ام از ماتریس $\pi(q, n, d, k)$ برابر یک است اگر و تنها اگر $\alpha \prec \tau$ باشد. به‌وضوح وزن سطر و ستون یک ماتریس $\pi(q, n, d, k)$ به‌ترتیب $\binom{n-d}{k-d}q^{k-d}$ و $\binom{k}{d}$ است. در جدول (۲.۵) یک $\pi(2, 4, 2, 3)$ را داریم.

جدول ۲.۵: $\pi(2, 4, 2, 3)$

$\pi(2, 4, 2, 3)$	۰۱۱۱	۰۱۱۲	۰۱۲۲	۰۲۱۲	۰۲۲۲	۱۰۱۱	۱۰۱۲	۱۰۲۱	۱۰۲۲	۱۱۰۱	...	۲۲۲۰
۰۰۱۱	۱	۰	۰	۰	۰	۱	۰	۰	۰	۰		۰
۰۰۱۲	۰	۱	۰	۱	۰	۰	۱	۰	۰	۰	...	۰
۰۰۲۱	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰		۰
۰۰۲۲	۰	۰	۱	۰	۱	۰	۰	۰	۱	۰		۰
۰۱۰۱	۱	۰	۰	۰	۰	۰	۰	۰	۰	۱		۰
۰۱۰۲	۰	۱	۱	۰	۰	۰	۰	۰	۰	۰	...	۰
۰۱۱۰	۱	۱	۰	۰	۰	۰	۰	۰	۰	۰		۰
۰۱۲۰	۰	۰	۱	۰	۰	۰	۰	۰	۰	۰		۰
⋮				⋮				⋮				
۲۱۰۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰

فرض می‌کنیم q توانی از یک عدد اول باشد. فضای برداری n بعدی روی F_q را با F_q^n نمایش می‌دهیم.

برای $1 \leq k \leq n$ تعداد زیرفضاهای k -بعدی روی F_q^n برابر ضرب گاوسی $\binom{n}{k}_q$ است و خانواده‌ی زیرفضاهای

k -بعدی روی F_q^n را با $\binom{[n]}{k}_q$ نمایش می‌دهیم. از [۳۷] داریم:

$$\binom{n}{k}_q = \prod_{i=0}^{k-1} (q^{n-i} - 1)(q^{k-i} - 1)^{-1}.$$

تعریف ۳.۵.۵. [۱۶] فرض می‌کنیم $1 \leq d \leq k \leq n$ و q برابر توانی از یک عدد اول باشد. $(0, 1)$ -ماتریس

$\gamma(n, d, k)$ را در نظر می‌گیریم، به‌طوری که سطرهای آن را اعضای $\binom{[n]}{d}_q$ و ستون‌های آن را اعضای $\binom{[n]}{k}_q$ تشکیل

می‌دهند.

برای $D \in \binom{[n]}{d}_q$ و $K \in \binom{[n]}{k}_q$ ، (D, K) -امین درایه از ماتریس برابر یک است اگر و تنها اگر D زیرمجموعه‌ی

(یا زیر فضای) K باشد. به‌وضوح وزن سطر و ستون یک ماتریس $\gamma(n, d, k)$ به‌ترتیب $\binom{n-d}{k-d}_q$ و $\binom{k}{d}_q$ است.

گزاره ۴.۵.۵. [۱۶] فرض می‌کنیم $1 \leq q$ ، $1 \leq s \leq d \leq k \leq n$ و $e = \binom{k-s}{d-s} - 1$ است.

الف. $(0, 1)$ -ماتریس $\delta(n, d, k)$ ، $(s, e+1)$ -مجزا است.

ب. $(0, 1)$ -ماتریس $\pi(q, n, d, k)$ ، $(s, e+1)$ -مجزا است.

اثبات. الف. فرض می‌کنیم $S = \{K_0, K_1, \dots, K_s\}$ خانواده‌ای $(s+1)$ -عضوی از k -مجموعه‌های دلخواه

ستون‌های $\delta(n, d, k)$ باشد. کافی است نشان دهیم که حداقل $\binom{k-s}{d-s}$ مجموعه‌ی d -عضوی وجود دارد که

زیرمجموعه‌ی K_0 هستند. ولی زیرمجموعه‌ی هیچ‌یک از K_i ها، $1 \leq i \leq s$ ، نیستند.

از آنجا که برای $1 \leq i \leq s$ ، $K_0 \neq K_i$ و $|K_0| = |K_i|$ لذا :

$$\exists x_1 \in K_0 \setminus K_1$$

$$\exists x_2 \in K_0 \setminus K_2$$

$\vdots$

$$\exists x_i \in K_0 \setminus K_i$$

$\vdots$

$$\exists x_s \in K_0 \setminus K_s.$$

فرض می‌کنیم $B = \{x_1, x_2, \dots, x_s\}$ و $|B| = s' \leq s$.

به‌وضوح تعداد زیرمجموعه‌های d -عضوی از K_0 که B زیرمجموعه‌ی آن‌ها است همان سطرهایی هستند

که توسط K_0 پوشیده می‌شوند ولی توسط هیچ یک از K_i ها، $1 \leq i \leq s$ ، پوشیده نمی‌شوند. تعداد چنین

زیرمجموعه‌هایی برابر است با:

$$\binom{k-s'}{d-s'} = \binom{k-s'}{k-d} \geq \binom{k-s}{k-d} = \binom{k-s}{d-s}.$$

اثبات. ب. مشابه قسمت (الف).

□

گزاره ۵.۵.۵.۵. [۱۶] برای $1 \leq d \leq k \leq n$ و $1 \leq r < k$ فرض می‌کنیم Γ خانواده‌ای از k -مجموعه‌ها روی

$[n]$ با مینیمم فاصله همینگ $2r$ باشد. قرار می‌دهیم:

$$p = \lceil \left(\binom{k}{d} - \binom{k-r}{d} \right) \left(\binom{k-r}{d} - \binom{k-2r}{d} \right)^{-1} \rceil,$$

و همچنین برای $1 \leq s \leq p$ قرار می‌دهیم:

$$e = \binom{k}{d} - \binom{k-r}{d} - (s-1) \left(\binom{k-r}{d} - \binom{k-2r}{d} \right) - 1,$$

سپس برای $1 \leq s \leq p$ ماتریس $\delta(n, d, k, \Gamma, r)$ ، $(s, e+1)$ -مجزا است.

اثبات. فرض می‌کنیم $S = \{K_0, K_1, \dots, K_s\}$ خانواده‌ای $(s+1)$ -عضوی از k -مجموعه‌های دلخواه ستون‌های $\delta(n, d, k, \Gamma, r)$ باشد. کافی است نشان دهیم که حداقل $e+1$ مجموعه‌ی d -عضوی وجود دارد که زیرمجموعه‌ی K_0 هستند ولی زیرمجموعه‌ی هیچ‌یک از K_i ها، $1 \leq i \leq s$ ، نیستند.

بدون اینکه به کلیت قضیه خللی وارد شود برای $1 \leq i \leq s$ فرض می‌کنیم $|K_0 \cap K_i| = k - r$. زیرا در این صورت خانواده $\{K_1, K_2, \dots, K_s\}$ حداکثر تعداد ممکن از d -مجموعه‌هایی که در K_0 هستند را می‌پوشاند. K_1 ، d -مجموعه از K_0 را می‌پوشاند. از آنجا که $|K_0 \cap K_i \cap K_{i+1}| \geq k - 2r$ لذا $\binom{k-r}{d} - \binom{k-2r}{d}$ تعداد d -مجموعه‌هایی است که به وسیله K_{i+1} پوشیده می‌شود اما به وسیله K_i پوشیده نمی‌شود. لذا نتیجه برقرار است. $\square$

گزاره ۶.۵.۵. [۱۶] فرض می‌کنیم q برابر توانی از یک عدد اول باشد. برای $1 \leq d \leq k \leq n$ قرار می‌دهیم:

$$p = \lceil \left(\binom{k}{d}_q - \binom{k-1}{d}_q \right) \left(\binom{k-1}{d}_q - \binom{k-2}{d}_q \right)^{-1} \rceil,$$

و همچنین برای $1 \leq s \leq p$ قرار می‌دهیم:

$$e = \binom{k}{d}_q - \binom{k-1}{d}_q - (s-1) \left(\binom{k-1}{d}_q - \binom{k-2}{d}_q \right) - 1,$$

سپس برای $1 \leq s \leq p$ ماتریس $\gamma(n, d, k)$ ، $(s, e+1)$ -مجزا است.

اثبات. فرض می‌کنیم $S = \{K_0, K_1, \dots, K_s\}$ خانواده‌ای از $(s+1)$ زیرفضای k -بعدی دلخواه از ستون‌های $\gamma(n, d, k)$ باشد. کافی است نشان دهیم که حداقل $e+1$ زیرفضای d -بعدی وجود دارد که زیر فضای K_0 هستند ولی زیرفضای هیچ‌یک از K_i ها، $1 \leq i \leq s$ ، نیستند.

بدون اینکه به کلیت قضیه خللی وارد شود برای $1 \leq i \leq s$ فرض می‌کنیم بعد $K_0 \cap K_i$ برابر $k-1$ باشد. زیرا در این صورت خانواده $\{K_1, K_2, \dots, K_s\}$ حداکثر تعداد زیرفضاهای d -بعدی که در K_0 هستند را می‌پوشانند. $K_1, \binom{k-1}{d}$ ، زیرفضای d -بعدی از K_0 را می‌پوشاند. از آنجایی که فصل مشترک هر دو زیرفضای $(k-1)$ -بعدی از یک زیرفضای k -بعدی، زیرفضای $(k-2)$ -بعدی است لذا تعداد زیرفضاهای d -بعدی از K_0 که به وسیله K_{i+1} پوشیده می‌شود ولی به وسیله K_i پوشیده نمی‌شود برابر $\binom{k-2}{d}_q - \binom{k-1}{d}_q$ است، لذا حکم برقرار است. $\square$

از گزاره فوق نتیجه زیر را داریم.

نتیجه ۲.۵.۵. [۱۶] فرض می‌کنیم q برابر توانی از یک عدد اول باشد. برای $1 \leq s \leq q$ قرار می‌دهیم:

$$e = q^{k-1} - (s-1)q^{k-2} - 1,$$

سپس برای $1 \leq s \leq q$ ماتریس $\gamma(n, 1, k)$ ، $(s, e+1)$ -مجزا است.

فصل ۶

چه موقع برای آزمون گروهی ایستا، آزمون انفرادی بهینه است؟

۱.۶ مقدمه

در آزمون گروهی هدف این است که با فرض وجود d بیمار در میان n نفر با حداقل آزمایش ممکن بتوانیم تمام بیماران را شناسایی کنیم. آنچه که در این فصل مورد مطالعه قرار می‌گیرد این است که با فرض معلوم بودن حداکثر مقدار d ، برای چه مقادیری از n در آزمون گروهی ایستا، آزمون انفرادی بهینه است. یک الگوریتم بدیهی که از ایده تست گروهی استفاده نمی‌کند، این است که همه افراد را به صورت انفرادی آزمایش می‌کند لذا برای n نفر نیاز به n آزمایش است. فرض می‌کنیم که برای ثابت d ، $N(d)$ بزرگترین مقدار n باشد که برای آن آزمون انفرادی بهینه است. در این فصل اثبات می‌کنیم که با در نظر گرفتن شرایطی در الگوریتم ایستا $N(d) = (d + 1)^2$ و همچنین اثبات می‌کنیم که برای $d = 1, 2, 3, 4$ ، $N(d) = (d + 1)^2$ است.

۲.۶ یک شرط لازم

فرض می‌کنیم $t(d, n)$ کمترین مقدار t ای باشد که برای آن یک ماتریس $M_{t \times n}$ ، d -مجزا داریم.

لم ۱.۲.۶. [۲۰] اگر $M_{t \times n}$ یک ماتریس d -مجزا باشد که $t < n$ آن‌گاه $N(d) \leq t$.

اثبات. می‌دانیم که $t(d, n)$ غیرکاهشی است و کمترین تعداد آزمون‌ها را برای یک ماتریس d -مجزا با n سطر نشان می‌دهد. بنابراین وجود یک ماتریس $M_{t \times n}$ ، d -مجزا که $t < n$ بر این امر دلالت می‌کند که می‌توان d بیمار را از میان $t + 1$ نفر با t آزمون مشخص کرد و این معادل است که بگوییم برای $n = t + 1$ بهترین الگوریتم، الگوریتم بدیهی نیست و لذا $N(d) \leq t$. $\square$

ضرب داخلی دو ستون C و C' را با $\lambda_{CC'}$ نمایش می‌دهیم و برای یک $(0, 1)$ -ماتریس M ، قرار می‌دهیم $\bar{\lambda} = \max \lambda_{CC'}$ که در آن ماکسیمم روی همه‌ی ستون‌های C و C' از M گرفته شده است. هر $(0, 1)$ -ماتریس دارای وزن w است اگر هر ستون آن دارای وزن w باشد.

لم ۲.۲.۶. [۳۰] هر ماتریسی با وزن w ، $\left(\lceil \frac{w}{\bar{\lambda}} \rceil - 1\right)$ -مجزا است.

اثبات. فرض می‌کنیم ماتریس M دارای وزن w باشد ولی برای $d, d = \left(\lceil \frac{w}{\bar{\lambda}} \rceil - 1\right)$ -مجزا نباشد. لذا d ستون $C_{i_1}, C_{i_2}, \dots, C_{i_d}$ و ستون C متمایز از این d ستون چنان وجود دارند که C زیرمجموعه‌ی اجتماع این d ستون است. لذا داریم:

$$w \leq \sum_{j=1}^d C C_{i_j} \leq \sum_{j=1}^d \bar{\lambda} = d \bar{\lambda}$$

$$\Rightarrow \frac{w}{\bar{\lambda}} + 1 \leq \lceil \frac{w}{\bar{\lambda}} \rceil$$

$$\Rightarrow \lceil \frac{w}{\bar{\lambda}} \rceil - \frac{w}{\bar{\lambda}} \geq 1.$$

$\square$

که این تناقض است و لذا حکم برقرار است.

نتیجه ۳.۲.۶. [۲۰] اگر $\bar{\lambda} = 1$ آن گاه یک ماتریس با وزن w ، $(w - 1)$ -مجزا است.

اثبات. با قرار دادن $\bar{\lambda} = 1$ در لم (۲.۲.۶) حکم برقرار است. $\square$

نتیجه مهم این بخش قضیه‌ی زیر است.

قضیه ۴.۲.۶. [۲۰] اگر $d + 1$ برابر توانی از یک عدد اول باشد، آن گاه $N(d) \leq (d + 1)^2$.

اثبات. با استفاده از لم (۱.۲.۶) و نتیجه (۳.۲.۶) کافی است تا یک ماتریس، $M_{(d+1)^2 \times n}$ با وزن $d + 1$ و $\bar{\lambda} = 1$ که $(d + 1)^2 < n$ بسازیم. از آنجا که $d + 1$ برابر توانی از یک عدد اول است لذا d مربع لاتین دوبه‌دو متعامد $L_1, L_2, \dots, L_d$ از مرتبه $d + 1$ وجود دارد [۲۶]. هر مربع لاتین $(d + 1)^2$ درایه دارد که شماره درایه‌های هر سطر و ستون آن زیرمجموعه‌ای از $\{0, 1, \dots, (d + 1)^2 - 1\}$ است. حال با اعمال جایگشت مناسب روی درایه‌های هر مربع لاتین می‌توانیم فرض کنیم که در سطر اول هر d مربع لاتین به ترتیب اعداد $d, \dots, 1, 0$ قرار دارند. دقت کنید که در این حالت درایه‌ی (i, j) ، $0 \leq i, j \leq d$ ، را معادل عدد $i(d + 1) + j$ در نظر می‌گیریم. در ماتریس $M_{(d+1)^2 \times n}$ ، $(d + 1)^2$ سطر به شماره‌های $0, 1, \dots, (d + 1)^2 - 1$ وجود دارد و هر ستون از ماتریس M زیرمجموعه‌ای از سطرها را مشخص می‌کند. یعنی برای هر $0 \leq k \leq d$ مجموعه‌ی شماره‌هایی از L_1 را که دارای درایه‌ی k هستند را به عنوان ستونی از M قرار می‌دهیم. لذا هر مربع لاتین می‌تواند $d + 1$ ستون از ماتریس M را بسازد. با توجه به تعریف مربع لاتین، هر یک از اعداد مجموعه‌ی $\{0, 1, \dots, d\}$ در هر سطر و ستون از L_1 دقیقاً یکبار ظاهر می‌شود. لذا وزن ستون‌های تولید شده از L_1 برابر $d + 1$ است و همچنین به ازای هر دو ستون C و C' از مجموعه ستون‌های تولید شده از L_1 ، $\lambda_{CC'} = 0$ است.

حال به سراغ مربع لاتین L_2 می‌رویم و $d + 1$ ستون دیگر را ایجاد می‌کنیم و این روند را برای سایر مربع لاتین‌های دیگر انجام می‌دهیم.

با توجه به تعریف مربع لاتین متعامد، اگر دو مربع لاتین متعامد را روی هم قرار دهیم آن گاه هر زوج (a, b) ، $0 \leq a, b \leq d$ دقیقاً یکبار ظاهر می‌شود. اگر ستون دلخواه C تولید شده از L_i و ستون دلخواه C' تولید شده

از L_j را در نظر بگیریم $\lambda_{CC'} = 1$ است. بنابراین با استفاده از d مربع لاتین متعامد می‌توان $d(d+1)$ ستون از ماتریس M را ساخت به طوری که $\bar{\lambda} = 1$ و وزن هر ستون $d+1$ است.

حال ماتریس $S_{(d+1) \times (d+1)}$ را در نظر می‌گیریم که در آن درایه‌ی (i, j) با مقدار $i(d+1) + j$ پر شده است. واضح است هر سطر و ستون از S نشان‌دهنده‌ی زیرمجموعه‌ای از $\{0, 1, \dots, (d+1)^2 - 1\}$ است. لذا می‌توان آن‌ها را به عنوان ستون‌هایی به ماتریس M اضافه کرد. به ازای هر دو ستون تولید شده از ستون‌های (سطرهای) ماتریس S ، چون این ستون‌ها (سطرها) هیچ اشتراکی باهم ندارند لذا $\lambda_{CC'} = 0$. به ازای هر ستون C ، تولید شده از ستونی از ماتریس S ، و ستون R ، تولید شده از سطری از ماتریس S ، چون دارای دقیقاً یک اشتراک هستند $\lambda_{CR} = 1$ است. همچنین از آنجا که در هر سطر و ستون از یک مربع لاتین هر درایه دقیقاً یک بار ظاهر می‌شود لذا برای هر ستون جدید C از M و ستون قدیمی آن مانند C' ، $\lambda_{CC'} = 1$. بنابراین برای ماتریس M که به این طریق ساختیم $\bar{\lambda} = 1$ و وزن هر ستون نیز $d+1$ است. لذا:

$$n = d(d+1) + 2(d+1) = (d+1)(d+2) > (d+1)^2.$$

پس یک ماتریس $M_{(d+1)^2 \times n}$ با $n > (d+1)^2$ و $\bar{\lambda} = 1$ و همچنین وزن $d+1$ ساختیم. لذا طبق نتیجه (۳.۲.۶) ماتریس M ، d -مجزا بوده و طبق لم (۱.۲.۶)، $N(d) \leq (d+1)^2$ است. $\square$

۳.۶ شرط لازم و کافی تحت $\bar{\lambda} = 1$

ساختن ماتریس d -مجزا با ساده‌ترین و متداول‌ترین روش‌های ارائه شده به وسیله نتیجه (۳.۲.۶) کار مشکلی است. در این بخش مسأله ارائه شده در عنوان فصل را تحت $\bar{\lambda} = 1$ مورد بررسی قرار می‌دهیم.

(۰، ۱)-ماتریس $M_{t \times n}$ را در نظر می‌گیریم و فرض می‌کنیم گراف G گرافی باشد که رأس‌های آن را سطرهای ماتریس تشکیل می‌دهند و بین دو رأس یال وجود دارد اگر و فقط اگر ضرب داخلی سطرهای متناظر آن دو رأس غیرصفر باشد. $d_G(v)$ را درجه‌ی رأس v در گراف G تعریف می‌کنیم.

لم ۱.۳.۶. [۲۰] اگر ماتریس $M_{t \times n}$ دارای وزن $d + 1$ و $\bar{\lambda} = 1$ باشد، آن گاه گراف $G(M)$ شامل n زیرگراف کامل K_{d+1} یال مجزا است.

اثبات. از آنجا که وزن هر ستون $d + 1$ است لذا هر یک از n ستون ماتریس یک گراف کامل از درجه d را تولید می کنند. چون $\bar{\lambda} = 1$ یعنی n زیرگراف کامل یال مجزا داریم. در غیر این صورت حداقل دو زیرگراف کامل وجود دارد که دارای حداقل یک یال مشترک $a_i a_j$ هستند. حال دو ستون متناظر با این دو زیرگراف را در نظر می گیریم که در سطرها a_i و a_j درایه ی یک دارند. لذا داریم $\bar{\lambda} > 1$ که این تناقض با فرض مسئله است. $\square$

لم ۲.۳.۶. [۶۰] هر ماتریس d - مجزا با $\bar{\lambda} = 1$ می تواند به یک ماتریس d - مجزای دیگر با وزن حداکثر $d + 1$ کاهش پیدا کند.

اثبات. ماتریس $M_{t \times n}$ ، d - مجزا با $\bar{\lambda} = 1$ را در نظر می گیریم. بدون اینکه به کلیت مسئله خللی وارد شود ستون های ماتریس را به گونه ای جابه جا می کنیم که ستون های دارای وزن بیشتر از $d + 1$ در ستون های اولیه ماتریس قرار بگیرند. فرض می کنیم ستون C_1 دارای وزن بیشتر از $d + 1$ باشد، یعنی $w(C_1) \geq d + 2$. نشان می دهیم که اگر یکی از درایه های یک واقع در این ستون را به صفر تبدیل کنیم ماتریس M_1 حاصل همچنان یک ماتریس d - مجزا باقی می ماند. در این صورت این روند را تا جایی ادامه می دهیم که وزن ستون C_1 ، $d + 1$ شود، سپس به سراغ ستون دیگر می رویم.

یکی از درایه های یک واقع در ستون C_1 را به صفر تبدیل می کنیم تا ماتریس M_1 ایجاد شود. به ازای هر زیرمجموعه ی D ، $|D| = d$ ، از ستون های ماتریس M_1 که شامل ستون C_1 باشد نمی توان ستونی را خارج از این d ستون پیدا کرد که زیرمجموعه ی اجتماع این d ستون باشد زیرا ماتریس M ، d - مجزا بوده است.

در حالت بعد فرض می کنیم زیرمجموعه ی D شامل ستون اول نباشد. در این صورت بنابر d - مجزا بودن ماتریس M ، هر ستون ماتریس M_1 بجز احتمالاً C_1 زیرمجموعه ی اجتماع این d ستون نیست. حال نشان می دهیم که C_1 نیز زیرمجموعه ی اجتماع این d ستون نیست. فرض می کنیم ستون C_1 در ماتریس M_1 زیرمجموعه ی اجتماع

ستون‌های واقع در D باشد. لذا:

$$\sum_{j \in D} C_1 C_j \geq d + 1,$$

از طرفی چون $C_1 C_j \leq 1$ لذا داریم:

$$d + 1 \leq \sum_{j \in D} C_1 C_j \leq |D| = d,$$

که این تناقض است. لذا C_1 زیرمجموعه‌ی اجتماع d ستون واقع در D نیست پس ماتریس M_1, d -مجزا است. $\square$

ستون a_i در ماتریس M را منفرد^۱ گوئیم هرگاه حداقل یک سطر موجود باشد که تنها در ستون a_i دارای درایه‌ی یک باشد. بعبارت دیگر حداقل یک گروه تک عضوی شامل a_i وجود داشته باشد.

$n(d, t)$ را بزرگترین مقدار n تعریف می‌کنیم به‌طوری‌که برای آن یک ماتریس d -مجزای $M_{t \times n}$ داریم.

لم ۳.۳.۶. [۲۵] اگر ماتریس d -مجزای $M_{t \times n}$ یک ستون منفرد داشته باشد، آن‌گاه:

$$n \leq 1 + n(d, t - 1).$$

اثبات. با حذف ستون منفرد و سطر متناظر با آن در ماتریس $M_{t \times n}$ ، ماتریس حاصل $M_{(t-1) \times (n-1)}$ همچنان

d -مجزا باقی می‌ماند. بنابراین $n - 1 \leq n(d, t - 1)$. $\square$

لم ۴.۳.۶. [۴۸] اگر ماتریس $d, M_{t \times n}$ -مجزا باشد، آن‌گاه هرستون با وزن حداکثر d منفرد است.

اثبات. فرض می‌کنیم ماتریس $d, M_{t \times n}$ -مجزا باشد و همچنین ستون C_1 دارای وزن $w \leq d$ باشد. بدون اینکه به کلیت مسأله خللی وارد شود می‌توان فرض کرد که این w درایه‌ی ناصفر در ستون C_1 در w سطر اول باشند. حال اگر حکم درست نباشد برای هر سطر $i, 1 \leq i \leq w$ ، لااقل یک ستون بجز C_1 وجود دارد که دارای درایه‌ی یک در این سطر است. فرض می‌کنیم سطر i -ام در ستون $x(i)$ ناصفر است که $2 \leq x(i) \leq n$ ، یعنی درایه‌ی

^۱isolated

$(i, x(i))$ برای $1 \leq i \leq w$ ناصفر است. ستون‌های واقع در $W = \{x(1), x(2), \dots, x(w)\}$ که $|W| \leq w$ را در نظر گرفته و $d - |W|$ ستون دلخواه دیگر از باقیمانده ستون‌ها را به آن اضافه می‌کنیم. بدیهی است که ستون C_1 زیرمجموعه‌ی اجتماع این d ستون است و این با d -مجزا بودن M در تناقض است. $\square$

لم ۵.۳.۶. [۲۰] اگر در ماتریس d -مجزای $M_{t \times n}$ با $\bar{\lambda} = 1$ ، وزن هر ستون حداقل $d + 1$ باشد، آنگاه داریم:

$$n \leq \left\lfloor \frac{t(t-1-r)}{d(d+1)} \right\rfloor,$$

که در آن r باقیمانده $t - 1$ به هنگ d است.

اثبات. درایه‌های یک ستون‌هایی از ماتریس $M_{t \times n}$ را که دارای وزن $k \geq d + 2$ است را به صفر تبدیل می‌کنیم تا جایی که وزن هر ستون برابر $d + 1$ شود. بنابر لم (۲.۳.۶) ماتریس حاصل همچنان d -مجزا باقی می‌ماند. این ماتریس جدید را $M_{t \times n}^*$ می‌نامیم. حال گراف $G(M^*)$ را در نظر می‌گیریم. طبق لم (۱.۳.۶) این گراف شامل n گراف کامل K_{d+1} مجزای K_{d+1} است، بنابراین $d|d_G(v)$. از آنجا که $G(M^*)$ دارای t رأس است لذا برای هر رأس v ، $d_G(v) \leq t - 1$ است. اما $d|d_G(v)$ و لذا باید $d_G(v) \leq t - 1 - r$ باشد. با توجه به اینکه تعداد یال‌ها در گراف کامل K_{d+1} برابر $\binom{d+1}{2}$ است بنابراین تعداد یال‌ها در گراف G برابر است با:

$$n \times \binom{d+1}{2} = \sum \frac{d_G(v)}{2} \leq \frac{t[(t-1)-r]}{2}.$$

$\square$

که به‌سادگی حکم از آن نتیجه می‌شود.

نتیجه مهم این بخش قضیه‌ی زیر است.

قضیه ۶.۳.۶. [۲۰] اگر $\bar{\lambda} = 1$ و $d + 1$ برابر توانی از یک عدد اول باشد، آنگاه $N(d) = (d + 1)^2$.

اثبات. در قضیه (۴.۲.۶) نشان دادیم $N(d) \leq (d + 1)^2$ ، حال کافی است تا نشان دهیم $N(d) \geq (d + 1)^2$. فرض می‌کنیم چنین نباشد و $N(d) \leq (d + 1)^2 - 1$ و این یعنی برای $n = (d + 1)^2$ آزمون انفرادی بهینه نیست. لذا یک ماتریس d -مجزا مانند $M_{t \times n}$ داریم که $t < n$ و همچنین با اضافه کردن سطر در صورت نیاز می‌توانیم فرض کنیم $t = (d + 1)^2 - 1$. فرض می‌کنیم ماتریس $M_{[(d+1)^2-1] \times (d+1)^2}$ ، d -مجزا با $\bar{\lambda} = 1$ باشد. حال با استفاده از لم (۲.۳.۶) می‌توان فرض کرد که هر ستون ماتریس M دارای وزن حداکثر $d + 1$ است. دو حالت زیر را داریم:

حالت اول: هر ستون از ماتریس M دارای وزن $d + 1$ باشد. از آنجا که:

$$t = (d + 1)^2 - 1 \Rightarrow t - 1 = (d + 1)^2 - 2 = d(d + 1) + d - 1 \Rightarrow t - 1 - r = d(d + 1),$$

با استفاده از لم (۵.۳.۶) داریم:

$$n \leq \frac{t(t-1-r)}{d(d+1)} = \frac{t(d+1)d}{d(d+1)} = t.$$

ولی این با فرض $t < n$ در تناقض است.

حالت دوم: فرض می‌کنیم ستون‌هایی با وزن حداکثر d در ماتریس M وجود داشته باشد و همچنین فرض می‌کنیم S مجموعه ستون‌هایی از ماتریس M با وزن حداکثر d باشد. با توجه به لم (۴.۳.۶) هر ستون $C \in S$ منفرد است. فرض می‌کنیم $r(C)$ تنها سطر برخورد کننده به ستون C باشد. حال ماتریس M' به دست آمده از ماتریس M با حذف C و $\{r(C), C \in S\}$ را در نظر می‌گیریم. در این صورت M' ماتریسی است از مرتبه $(t - |S|) \times (n - |S|)$ با وزن $d + 1$ و $\bar{\lambda} = 1$ و از آنجا که:

$$t - |S| - 1 = d(d + 1) + d - 1 - |S| \Rightarrow t - 1 - |S| - r = d(d + 1).$$

با توجه به لم (۵.۳.۶) داریم:

$$(n - |S|) \leq \lfloor \frac{(t-|S|)d(d+1)}{d(d+1)} \rfloor,$$

□

لذا $n \leq t$ که در این حالت نیز تناقض است.

۴.۶ بدست آوردن $N(d)$ برای مقادیر کوچک d

باسالایگو در [۱۴] لم زیر را اثبات کرد.

لم ۱.۴.۶. اگر $M_{t_* \times n}$ یک ماتریس d -مجزا و C ستونی از ماتریس M با وزن w باشد، آنگاه

$$t_* \geq w + t(d-1, n-1).$$

اثبات. ستون C و سطرهاى برخوردکننده به آن را حذف می‌کنیم. در این صورت یک ماتریس $(d-1)$ -مجزا با $n-1$ ستون و $t_* - w$ سطر به وجود می‌آید. از آنجا که $t(d-1, n-1)$ کمترین تعداد آزمون‌هایی است که برای آن یک ماتریس $M_{t \times (n-1)}$ ، $(d-1)$ -مجزا داریم لذا:

$$t_* - w \geq t(d-1, n-1) \quad \Rightarrow \quad t_* \geq w + t(d-1, n-1).$$

□

لم معروف اسپنسر را به صورت زیر داریم.

$$\text{لم ۲.۴.۶. } n(1, t) = \binom{t}{\lfloor \frac{t}{2} \rfloor} \quad [۳۳]$$

$$\text{قضیه ۳.۴.۶. } N(1) = 4 \quad [۲۵]$$

اثبات. $d+1 = 2$ عدد اولی است لذا بنابر قضیه (۴.۲.۶)، $N(1) \leq 4$. بنابر لم (۲.۴.۶)، $n(1, 3) = 3$

است، یعنی بزرگترین مقدار n که برای آن یک ماتریس $M_{3 \times n}$ ، ۱-مجزا داریم برابر ۳ است.

اگر $n_0 < 4$ ، $N(1) = n_0$ ، آن گاه یک ماتریس $M_{t \times 4}$ ، ۱- مجزا داریم که $t < 4$. در صورت نیاز با اضافه کردن سطر می توان فرض کرد که $t = 3$ و لذا $n(1, 3) \geq 4$ که این تناقض است. پس می توان نتیجه گرفت که $N(1) \geq 4$.
لذا $N(1) = 4$. $\square$

لم ۴.۴.۶. $n(2, 7) = 7$ [۲۰].

اثبات. ماتریس $M_{7 \times n}$ ، ۲- مجزا را در نظر می گیریم. دو حالت زیر را داریم:

الف. اگر $\bar{\lambda} = 1$ باشد چون $3 = d + 1$ عدد اولی است با استفاده از قضیه (۶.۳.۶)، $N(2) = 9$ است. فرض می کنیم $n(2, 7) = k$ باشد، بعبارت دیگر بزرگترین مقدار n که برای آن یک ماتریس $M_{7 \times n}$ ، ۲- مجزا داریم برابر k است. اگر $k > 7$ لذا یک ماتریس $M_{7 \times 8}^*$ ، ۲- مجزا داریم. لذا با توجه به لم (۱.۲.۶)، $N(2) \leq 7$ که این در تناقض با $N(2) = 9$ است. پس $k \leq 7$ و از طرفی همواره داریم $n(2, 7) = k \geq 7$ (زیرا ماتریس $I_{t \times t}$ یک ماتریس d - مجزا است). در نتیجه $n(2, 7) = 7$.

ب. در این حالت فرض می کنیم دو ستون C و C' موجود است که $\lambda_{CC'} > 1$. در این صورت ستون C یا منفرد است و یا اینکه دارای وزن حداقل چهار است زیرا:

چون $\lambda_{CC'} > 1$ لذا سطرهای r_1 و r_2 وجود دارند که ستون های C و C' در این سطرها دارای درایه ی یک هستند. بنابر خاصیت ۲- مجزایی سطری مانند r_3 وجود دارد که ستون C در آن سطر دارای درایه ی یک و ستون C' درایه صفر دارد و همچنین سطری مانند r_4 وجود دارد که ستون C' در آن سطر دارای درایه ی یک و ستون C درایه صفر دارد.

حال اگر در سطر r_3 تمام درایه ها بجز درایه ی ستون C صفر باشد ستون C منفرد است.

	C	C'		
r_1	۱	۱		
r_2	۱	۱		
r_3	۰	۱	۰	۰
r_4	۰	۱		

اگر ستونی مانند C'' وجود داشته باشد که در سطر r_3 دارای درایه‌ی یک باشد بنابر خاصیت ۲- مجزایی (چون $C' \not\subseteq C'' \cup C$) باید سطری مانند r_5 وجود داشته باشد که در آن سطر C دارای درایه‌ی یک باشد ولی C' و C'' دارای درایه‌ی صفر باشند. پس وزن ستون C حداقل چهار است.

	C	C'	C''
r_1	۱	۱	
r_2	۱	۱	
r_3	۱	۰	۱
r_4	۰	۱	
r_5	۱	۰	۰

پس در حالت ب مسأله در دو وضعیت بررسی می‌شود:

حالت اول: اگر ستون C منفرد باشد آن‌گاه با حذف ستون C و حذف سطری از ماتریس که تنها در ستون C دارای درایه‌ی یک است ماتریس ۲- مجزای، $M'_{6 \times (n-1)}$ را داریم. حال دو حالت زیر را در نظر می‌گیریم.

۱. اگر $\bar{\lambda} = 1$ باشد. فرض می‌کنیم $n(2, 6) = k > 6$ ، در این صورت یک ماتریس $M_{6 \times 7}$ ، ۲- مجزا داریم.

بنابر لم (۱.۲.۶)، داریم $N(2) \leq 6$ و بنابر لم (۱.۳.۱)، داریم $N(2) \geq \binom{4}{2} = 6$ پس $N(2) = 6$.

از طرفی طبق لم (۶.۳.۶) داریم $N(2) = 9$. حال که به تناقض رسیدیم نتیجه می‌شود که $k \leq 6$ و از

طرفی $n(2, 6) = k \geq 6$ پس $n(2, 6) = 6$. لذا $n(2, 7) \leq 1 + n(2, 6) \leq 1 + 6 = 7$.

۲. فرض می‌کنیم که در ماتریس $M'_{6 \times (n-1)}$ ، $\bar{\lambda} > 1$ است. لذا دو ستون C'' و C''' وجود دارند که $\lambda_{C''C'''} > 1$.

۱ است. لذا ستون C'' یا منفرد است و یا دارای وزن حداقل چهار است.

اگر ستون C'' در ماتریس $M'_{6 \times (n-1)}$ منفرد باشد با حذف ستون C'' و سطر مرتبط با آن یک ماتریس

$M_{5 \times (n-2)}^*$ داریم که ۲- مجزا است. حال اگر $n - 2 > 5$ آنگاه یک ماتریس $\bar{M}_{5 \times 6}$ ، ۲- مجزا داریم.

بنابر لم (۱.۲.۶)، $N(2) \leq 5$ و از طرفی بنابر لم (۱.۳.۱)، $N(2) \geq 6$ حال که به تناقض رسیدیم و لذا

باید $n - 2 \leq 5$ باشد و این یعنی $n(2, 7) \leq 7$ و از طرفی $n(2, 7) \geq 7$ پس $n(2, 7) = 7$.

فرض می‌کنیم ستون C'' در ماتریس $M'_{6 \times (n-1)}$ دارای وزن بیشتر یا مساوی چهار باشد. ستون C''' و

سطرهای برخورد کننده متناظر با آن را حذف می‌کنیم که در این صورت ماتریس $M''_{s \times (n-2)}$ حداکثر دو سطر دارد و ۱- مجزا است. با استفاده از لم (۲.۴.۶)، $n - 2 \leq 2$ یا $n \leq 4$.

حالت دوم: اگر C دارای وزن حداقل چهار باشد ستون C و سطرهای مرتبط با آن را حذف می‌کنیم. در این صورت ماتریس کاهش یافته دارای حداکثر سه سطر است و ۱- مجزا است که به وسیله لم (۲.۴.۶)، $n - 1 \leq 3$ یا $n \leq 4$. $\square$

قضیه ۵.۴.۶. $N(2) = 9$ [۲۰].

اثبات. از آنجایی که $d + 1 = 3$ یک عدد اول است، طبق قضیه (۴.۲.۶) داریم $N(2) \leq 9$. بنابراین کافی است تا نشان دهیم که $N(2) \geq 9$ یا $n(2, 8) = 8$. برای $\bar{\lambda} = 1$ طبق قضیه (۶.۳.۶) حکم برقرار است. فرض می‌کنیم که $M_{\lambda \times n}$ یک ماتریس ۲- مجزا باشد و دو ستون C و C' از ماتریس $M_{\lambda \times n}$ که $\lambda_{C, C'} > 1$ وجود داشته باشد. در این صورت دو حالت زیر را بررسی می‌کنیم:

۱. اگر C ستون منفرد باشد با حذف ستون C و سطر مرتبط با آن ماتریس کاهش یافته همچنان ۲- مجزا بوده و طبق دو لم (۳.۳.۶) و (۴.۴.۶) داریم:

$$n \leq 1 + n(2, 7) = 8,$$

و همچنین $n(2, 8) \geq 8$ ، لذا $n(2, 8) = 8$.

۲. حال اگر ستون C دارای وزن حداقل چهار باشد، با حذف ستون C و سطرهای برخورد کننده به آن ماتریس ۱- مجزای $M'_{s \times (n-1)}$ دارای حداکثر چهار سطر است که با استفاده از لم (۲.۴.۶) داریم:

$$s \leq 4 \Rightarrow n - 1 \leq 6 \Rightarrow n \leq 7,$$

$$\text{لذا } n(2, 8) \leq 7.$$

اثبات دیگری برای قسمت ۲. فرض کنید ستون C دارای حداقل وزن ۴ است. حال اگر $n \geq 9$ آن گاه یک ماتریس 9×8 ، 2 -مجزا داریم. بنابر لم (۱.۴.۶) داریم $t(1, 8) + 4 \geq 8$. از طرفی بنابر لم (۲.۴.۶)، باید $t(1, 8) \geq 5$ ولی این امکان پذیر نیست. پس $n \leq 8$ لذا $n(2, 8) = 8$.

□

$$\text{نتیجه ۶.۴.۶. [۲۰]} \quad N(3) \geq 13.$$

اثبات. با استفاده از لم (۴.۳.۶) و (۱.۴.۶) برای یک ماتریس 3 -مجزای $M_{t \times n}$ با حداقل یک ستون نامنفرد داریم:

$$t \geq (d+1) + t(d-1, n-1).$$

$$\text{از آنجا که } n(2, 8) = 8 \text{ لذا } t(2, 12) \geq 9. \text{ بنابراین برای } n \geq 13,$$

$$t \geq 4 + t(2, n-1) \geq 13.$$

□

$$\text{لذا برای } n \leq 13, t(3, n) = n \text{ و در نتیجه } N(3) \geq 13.$$

برای هر ستون C از یک ماتریس $M_{t \times n}$ ، $I(C)$ را خانواده‌ی زیرمجموعه‌های b -عضوی، $b \geq 2$ ، از مجموعه‌ی $\{1, 2, \dots, t\}$ در نظر می‌گیریم که برای هر $T \in I(C)$ ، لااقل یک ستون C' متفاوت از ستون C وجود دارد که $C' \cap C = T$.

به عنوان مثال اگر $I(C) = \{(1, 2), (1, 3, 4)\}$ یعنی اینکه حداقل ستونی مانند C_1 وجود دارد که فصل مشترک آن با ستون C در سطرهاى یک و دو است و حداقل یک ستون مانند C_2 وجود دارد که فصل مشترک آن با ستون C در سطرهاى یک، سه و چهار است.

حال فرض می‌کنیم M_1 ماتریسی $n_1 \times t$ ، 2 -مجزا با وزن $w = 4$ و بدون ستون منفرد باشد. در این صورت $\bar{\lambda} \leq 2$

است. زیرا اگر $\bar{\lambda} \geq 3$ باشد، یعنی حداقل دو ستون C_i و C_j از ماتریس M_1 وجود دارد که در سطرها r_1, r_2 و r_3 درایه‌ی یک دارند. از آنجایی که وزن ماتریس چهار است لذا فرض می‌کنیم ستون C_i در سطر دیگری مانند r_4 برابر یک است. از آنجاکه ماتریس ستون منفرد ندارد لذا ستون دیگری مانند C_k وجود دارد که در سطر r_4 برابر یک است. که این تناقض با ۲- مجزا بودن $(M_1)_{t \times n_1}$ است.

حال می‌خواهیم بعضی از درایه‌های ماتریس M_1 را به گونه‌ای حذف کنیم که وزن ماتریس از چهار به سه کاهش پیدا کند و همچنین خاصیت ۲- مجزا بودن بودن در ماتریس M_1^i ، ($i \geq 0$) باقی بماند (M_1^i ماتریس کاهش یافته بعد از i -امین بار عملیات حذف کردن روی ماتریس M_1 است). توجه کنید که حذف کردن هر درایه یک از ستون C روی $I(C')$ اثر می‌گذارد و ممکن است باعث تغییر آن شود. بنابراین بعد از هر مرحله حذف کردن باید $I(C)$ را برای ستون با وزن چهار C تشکیل دهیم. ستون C_j با وزن چهار را از ماتریس M_1^i که $i \geq 0$ و $I(C_j) \neq \phi$ در نظر می‌گیریم که برای حذف کردن دستورهای زیر را اجرا می‌کنیم. به راحتی می‌توان دید که $I(C_j)$ حداکثر دارای سه عضو است و در یکی از دو حالت زیر صدق می‌کند:

۱. اگر $I(C_j)$ در M_1^i به صورت زیر باشد:

$$I(C_j) \subseteq \{(x_j, y_j), (x_j, z_j), (x_j, v_j) \ ; \ x_j \neq y_j \neq v_j \in \{1, 2, \dots, t\}\}$$

درایه واقع در سطر x_j از ستون C_j را حذف می‌کنیم (در واقع به صفر تبدیل می‌کنیم). لذا این ستون کاهش یافته دارای وزن سه است و حاصل ضرب داخلی آن با سایر ستون‌های ماتریس حداکثر یک است. زیرا قبلاً ستون C_j با سایر ستون‌ها حداکثر در دو سطر یکسان دارای درایه یک بود که سطر x_j در تمام آن‌ها مشترک بوده، حال با حذف درایه‌ی یک در سطر x_j و ستون C_j به وضوح حاصل ضرب داخلی C_j با سایر ستون‌ها حداکثر یک است.

ماتریس M_1^i همچنان ۲- مجزا باقی می‌ماند. به وضوح اجتماع هیچ دو ستون که یکی از آن‌ها C_j است شامل هیچ ستون دیگری از ماتریس M_1^i نیست و همچنین اجتماع هیچ دو ستون دیگر غیر از C_j نیز شامل

هیچ ستون دیگری غیر از C_j نیست. اینکه آیا اجتماع دو ستون بجز C_j شامل C_j است یا خیر؟ جواب خیر است. زیرا می‌دانیم فصل مشترک ستون C_j با سایر ستون‌ها حداکثر یک سطر است و ستون C_j دارای وزن سه است لذا اجتماع هر دو ستون، شامل ستون C_j نیست. لذا M_1^i ماتریس ۲ - مجزا باقی می‌ماند.

۲. اگر $I(C_j)$ در M_1^i به صورت زیر باشد:

$$I(C_j) = \{(x_j, y_j), (x_j, z_j), (y_j, z_j)\},$$

در این حالت فعلاً هیچ کاری انجام نمی‌دهیم.

سرانجام به یک ماتریس $(M_1^*)_{t \times n_1}$ خواهیم رسید که برای هیچ یک از ستون‌های با وزن چهار آن حالت اول اتفاق نمی‌افتد. حال اگر در مورد ستون‌های با وزن چهار، حالت دوم رخ ندهد عملیات تمام است. ماتریس M_1^* را زیرماتریسی از M_1^* در نظر بگیرید که همه‌ی ستون‌های آن از نوع دوم هستند. این ماتریس دارای سه خاصیت زیر است:

الف. M_1^* دارای حداقل چهار ستون است که هر ستون دارای وزن چهار است. از آنجا که در M_1^* ستونی مانند C_j به صورت زیر وجود دارد:

$$I(C_j) = \{(x_j, y_j), (x_j, z_j), (y_j, z_j)\},$$

لذا M_1^* دارای حداقل چهار ستون است. فرض می‌کنیم C ستونی از M_1^* باشد که $C \cap C_j = \{x_j, y_j\}$ ولی C متعلق به M_1^* نباشد. از آنجا که هر ستون M_1 دارای وزن چهار بوده پس باید در مورد ستون C حالت اول اتفاق افتاده باشد یعنی یکی از حالات زیر:

$$I(C) \subseteq \{(x_j, y_j), (x_j, z_j), (x_j, v_j)\},$$

$$I(C) \subseteq \{(x_j, y_j), (z_j, y_j), (v_j, y_j)\},$$

لذا طبق این دو حالت باید x_j یا y_j از آن حذف شود که این تناقض است با $C \cap C_j = \{x_j, y_j\}$.
 بطورمشابه برای هر ستون C که $C \cap C_j = \{x_j, z_j\}$ یا $C \cap C_j = \{z_j, y_j\}$ نتیجه می‌شود که این ستون‌ها متعلق به ماتریس M_1^* است لذا این ماتریس حداقل چهار ستون دارد.

ب. اگر $C_j \in C(M_1^*)$ آن‌گاه $I(C_j) = \{(x_j, y_j), (x_j, z_j), (y_j, z_j)\}$. زیرا C_j از نوع دوم است.

پ. اگر $C_j \in M_1^*$ و ستون C_i از M_1 باشد که $|C_i \cap C_j| = 2$ باشد، آن‌گاه $C_i \in M_1^*$. فرض می‌کنیم $|C_i \cap C_j| = 2$. می‌دانیم که C_i در ماتریس M_1 دارای وزن چهار است. حال اگر C_i از نوع اول باشد لذا با حذف درایه یک آن طبق حالت اول به ماتریسی می‌رسیم که حاضرب داخلی C_i با سایر ستون‌ها یک است که با فرض $|C_i \cap C_j| = 2$ متناقض است. لذا از نوع دوم بوده و بنابراین $C_i \in M_1^*$.

پس M_1^* ماتریسی است که همه ستون‌های آن از نوع دوم هستند و همچنین دارای خاصیت‌های فوق است. لذا M_1' را می‌توان زیرماتریسی از M_1^* گرفت که نسبت به این سه خاصیت مینیمال باشد یعنی M_1' دارای هیچ زیرماتریسی با تعداد ستون‌های کمتر که در شرایط فوق صدق کند، نیست.

حال فرض می‌کنیم $(M_1'')_{q \times n'}$ زیرماتریسی است از M_1' که q سطر آن سطرهاى واقع در مجموعه‌ی زیر هستند:

$$\{x_i, y_i, z_i : C_i \in C(M_1') , I(C_i) = \{(x_i, y_i), (x_i, z_i), (y_i, z_i)\}\}.$$

لم ۷.۴.۶. [۲۰] برای ماتریس $(M_1'')_{q \times n'}$ داریم $n' \geq q$.

اثبات. فرض می‌کنیم $C \in M_1'$ لذا $I(C) = \{(x, y), (x, z), (y, z)\}$. پس هر سطر x, y, z در M_1'' باید حداقل سه درایه‌ی یک داشته باشند. برای مثال برای سطر x درایه‌ی یک، یک‌بار در C ظاهر می‌شود، یک‌بار در ستونی که با C در (x, y) مشترک است و یک‌بار با ستونی که با C در (x, z) مشترک است، ظاهر می‌شود. حال فرض می‌کنیم ستون C دارای چهارمین درایه‌ی یک باشد که در سطر v -ام از M_1'' ظاهر شده است. لذا طبق تعریف M_1'' باید $v \in \{x', y', z'\}$ ، بطوری‌که ستون C' ای در M_1' موجود باشد که $I(C') = \{(x', y'), (x', z'), (y', z')\}$.

بنابراین سطر v -ام حداقل چهار درایه‌ی یک دارد. فرض می‌کنیم k تعداد ستون‌هایی در M'' باشد که دارای وزن چهار است. حال با شمارش تعداد یک‌ها بوسیله سطرها و ستون‌ها داریم:

$$\mathfrak{Z}n'_1 + k = \text{حداکثر تعداد یک‌هایی که در } M'' \text{ ظاهر می‌شوند},$$

لذا:

$$\mathfrak{Z}n'_1 + k \geq \mathfrak{Z}q + k \Rightarrow n'_1 \geq q.$$

□

لم ۸.۴.۶. [۲۰] فرض کنید برای هر ستون C از M'_1 داریم $C = \{x, y, z, v\}$ و $I(C) = \{(x, y), (x, z), (y, z)\}$

در این صورت اگر M'_1 یک ماتریس ۲-مجزا باشد و همچنین $C' \cap \{x, y, z\} \neq \emptyset$ آن‌گاه $v \notin C'$.

اثبات. فرض می‌کنیم C' ستونی باشد که شامل v است. چون $C' \cap \{x, y, z\} \neq \emptyset$ بدون از دست دادن کلیت

مسأله فرض می‌کنیم $x \in C'$. حال اجتماع ستون C' به همراه ستون دیگری که فصل مشترک آن با ستون C ، (y, z)

است شامل ستون C است که این تناقض با ۲-مجزا بودن است. □

برای اثبات لم بعد تعریف زیر را ارائه می‌دهیم.

فرض می‌کنیم t, k, v و λ با فرض $t < k < v$ اعداد طبیعی باشند. منظور از یک t -طرح یا $(v, k, \lambda) - t$ طرح

یک زوج به صورت (X, B) است که در آن X مجموعه‌ای v عضوی از نقاط و B خانواده‌ای از زیرمجموعه‌های

k -عضوی از X بنام بلوک است بطوری که هر t نقطه از X در دقیقاً λ بلوک واقع شوند. یک $(v, 3, 1) - 2$

طرح، سیستم سه‌تایی اشتاینر یا به اختصار یک $STS(v)$ نامیده می‌شود.^۲

^۲Steiner triple system

$$\text{لم } ۰.۹۰۴۰۶. [۲۰] ۱۲ = n(۲, ۹).$$

اثبات. یک $(۱, ۳, ۹) - ۲$ طرح را در نظر می‌گیریم. حال ماتریس M را که نقاط X سطرها و بلوک‌ها ستون‌ها را تشکیل می‌دهند، در نظر می‌گیریم. از آنجایی که هر دو سطر از ماتریس در دقیقاً یک ستون یکسان ظاهر می‌شوند و لذا واضح است که ماتریس M ، ۲-مجزا است. از [۱۲] حداکثر تعداد بلوک‌های یک $(۱, k, v) - t$ برابر است با $\lfloor \lfloor \frac{v}{k} \rfloor \lfloor \frac{v-1}{k-1} \rfloor \dots \lfloor \frac{v-t+1}{k-t+1} \rfloor \dots \rfloor$. لذا در یک $(۱, ۳, ۹) - ۲$ طرح، خانواده B دارای حداکثر دوازده بلوک است که طرح معرفی شده در زیر یک $(۱, ۳, ۹) - ۲$ طرح با دوازده بلوک است.

$$B = \{\{۱, ۲, ۹\}, \{۳, ۴, ۹\}, \{۵, ۶, ۹\}, \{۷, ۸, ۹\}, \{۱, ۳, ۵\}, \{۲, ۴, ۶\}, \{۲, ۳, ۷\}, \{۱, ۴, ۸\},$$

$$\{۱, ۶, ۷\}, \{۲, ۵, ۸\}, \{۳, ۶, ۸\}, \{۴, ۵, ۷\}\}$$

با توجه به این طرح ماتریس M را به صورت زیر داریم:

$$M = \begin{pmatrix} ۱ & ۰ & ۰ & ۰ & ۱ & ۰ & ۰ & ۱ & ۱ & ۰ & ۰ & ۰ \\ ۱ & ۰ & ۰ & ۰ & ۰ & ۱ & ۱ & ۰ & ۰ & ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ & ۰ & ۱ & ۰ & ۱ & ۰ & ۰ & ۰ & ۱ & ۰ \\ ۰ & ۱ & ۰ & ۰ & ۰ & ۱ & ۰ & ۱ & ۰ & ۰ & ۰ & ۱ \\ ۰ & ۰ & ۱ & ۰ & ۱ & ۰ & ۰ & ۰ & ۰ & ۱ & ۰ & ۱ \\ ۰ & ۰ & ۱ & ۰ & ۰ & ۱ & ۰ & ۰ & ۱ & ۰ & ۱ & ۰ \\ ۰ & ۰ & ۰ & ۱ & ۰ & ۰ & ۱ & ۰ & ۱ & ۰ & ۰ & ۱ \\ ۰ & ۰ & ۰ & ۱ & ۰ & ۰ & ۰ & ۱ & ۰ & ۱ & ۱ & ۰ \\ ۱ & ۱ & ۱ & ۱ & ۰ & ۰ & ۰ & ۰ & ۰ & ۰ & ۰ & ۰ \end{pmatrix}$$

$$\text{بنابراین می‌توان نتیجه گرفت } n(۲, ۹) \geq ۱۲.$$

حال باید نشان دهیم که $n(۲, ۹) \leq ۱۲$. فرض کنید $M_{۹ \times n}$ یک ماتریس ۲-مجزا باشد که $n \geq ۱۳$. حال یک ستون دلخواه از M با وزن w را در نظر بگیرید. با حذف این ستون و سطرها متقاطع با این ستون یک ماتریس $M'_{(۹-w) \times (n-۱)}$ حاصل می‌شود که ۱-مجزا است. از آنجا که $۱۲ \geq n - ۱$ لذا بنا به لم (۲.۴.۶) باید $۹ - w \geq ۶$ و به عبارت دیگر $w \leq ۳$. بنابراین M یک ماتریس ۲-مجزا است که دارای وزن حداکثر ۳ است. اگر M دارای یک ستون با وزن کمتر از ۳ باشد آن‌گاه بنابه لم (۴.۳.۶)، این ستون منفرد است و لذا بنا به لم

(۳.۳.۶) باید $n \leq 1 + n(2, 8)$ اما بنا به قضیه‌ی (۵.۴.۶) می‌دانیم $n(2, 8) = 8$ و این یعنی $n \leq 9$ که با فرض $n \geq 13$ در تناقض است. بنابراین هر ستون از M دارای وزن ۳ است و M یک ماتریس ۲-مجزا و فاقد ستون منفرد است. با توجه به ۲-مجزا بودن به سادگی می‌توان دید هر دو زیرمجموعه دو عضوی از سطرها در حداکثر یک ستون واقع می‌شود. بنابراین $\frac{\binom{9}{2}}{3} \geq n$ و این یعنی $n \leq 12$ که با فرض متناقض است. $\square$

$$\text{لم } ۱۰۰۴۰۶. [۲۰] \leq n(2, 10).$$

اثبات. فرض می‌کنیم ماتریس $M_{10 \times n}$ یک ماتریس ۲-مجزا باشد. حال اگر ماتریس $M_{10 \times n}$ دارای حداقل یک ستون منفرد باشد آن‌گاه طبق لم (۹.۴.۶) و (۳.۳.۶) داریم:

$$n \leq 1 + n(2, 9) = 1 + 12 = 13 \Rightarrow n \leq 13.$$

اگر $\bar{\lambda} = 1$ و ماتریس M دارای هیچ ستون منفردی نباشد (لذا طبق عکس نقیض لم (۴.۳.۶) وزن هر ستون M باید $1 + d$ یا بیشتر از آن باشد) لذا با استفاده از لم (۵.۳.۶) از آنجا که $r \equiv t - 1 \equiv 10 - 1 \equiv 1 \pmod{2}$ داریم:

$$n \leq \left\lfloor \frac{t(t-1-r)}{(d+1)d} \right\rfloor = 13 \Rightarrow n \leq 13.$$

حال فرض می‌کنیم ماتریس $M_{10 \times n}$ دارای ستون منفرد نباشد و $\bar{\lambda} > 1$. فرض می‌کنیم C ستونی با ماکسیمم وزن w در ماتریس باشد، آن‌گاه C باید دارای وزن حداقل چهار باشد. زیرا اگر $w < 4$ با توجه به $\bar{\lambda} \geq 2$ و اینکه ماتریس دارای ستون منفرد نیست می‌توان با ۲-مجزا بودن به تناقض رسید.

۱. اگر وزن ستون C بیشتر از چهار باشد با استفاده از لم (۱.۴.۶) داریم:

$$t(2, n) \geq 5 + t(1, n-1),$$

$$10 \geq 5 + t(1, n-1) \Rightarrow t(1, n-1) \leq 5$$

و یا از لم (۲.۴.۶) داریم $n(2, 10) \leq 11$.

۲. فرض می‌کنیم وزن ستون C چهار است.

پس ماتریس M یک ماتریس ۲- مجزا با $\bar{\lambda} = 2$ است (زیرا اگر $\bar{\lambda} \geq 3$ با توجه به اینکه $w_C = 4$ و ماتریس ستون منفرد ندارد می‌توان با ۲- مجزا بودن به تناقض رسید). هر ستون از ماتریس M دارای وزن سه یا چهار است (زیرا اگر $w = 2$ با توجه به اینکه $\bar{\lambda} = 2$ می‌توان با ۲- مجزا بودن به تناقض رسید). فرض می‌کنیم $M = M_1 \cup M_2$ که M_1 یک ماتریس $10 \times n_1$ با وزن چهار و M_2 یک ماتریس $(n - n_1) \times 10$ با وزن سه باشد.

از آنجا که ماتریس M یک ماتریس ۲- مجزا با $\bar{\lambda} = 2$ است لذا M_1 و M_2 هر دو ۲- مجزا هستند و همچنین در ماتریس M_1 ، $\bar{\lambda} = 2$ و در ماتریس M_2 ، $\bar{\lambda} = 1$ است، زیرا اگر $\bar{\lambda} = 2$ باشد لذا دو ستون C_1 و C_2 در M_2 وجود دارند که $\lambda_{C_1 C_2} = 2$ حال اگر این دو ستون در ماتریس M بررسی شوند چون M بدون ستون منفرد است لذا ستونی مانند C' موجود است که اجتماع آن با C_2 شامل C_1 است که این تناقض با ۲- مجزا بودن است. به‌وضوح:

$$\forall C_1 \in M_1, C_2 \in M_2 \quad \lambda_{C_1 C_2} \leq 1$$

از طرفی ماتریس M_1 با $\bar{\lambda} = 2$ می‌تواند به یک ماتریس با $\bar{\lambda} = 1$ کاهش پیدا کند به‌طوری که ماتریس همچنان ۲- مجزا باقی بماند. روش قبل را استفاده می‌کنیم که در این صورت به ماتریس‌های M_1^* ، M_2^* ، M_1' و M_2' می‌رسیم که اگر ماتریس‌های M_1' و M_2' وجود نداشته باشند کار تمام است.

سطرها را به‌گونه‌ای جابه‌جا می‌کنیم به‌طوری که ستون C_1 برابر $\{1, 2, 3, 4\}$ باشد. می‌دانیم که $I(C_1)$ از نوع دوم است کامل بودن $I(C_1) = \{(1, 2), (1, 3), (2, 3)\}$ دلالت بر این دارد که زیرماتریس زیر باید در M_1' باشد.

توجه شود که برای کامل بودن $I(C_2)$ ، $I(C_3)$ و $I(C_4)$ نیاز به حداقل یک سطر بیشتر داریم. بنا به لم (۸.۴.۶) از آنجایی که هر یک از ستون‌های C_2 ، C_3 و C_4 با $\{x, y, z\}$ دارای اشتراک است لذا هیچ

یک از این ستون‌ها در سطر چهارم دارای درایه‌ی یک نیستند. حال فرض می‌کنیم ستون C_2 در سطر پنجم دارای درایه‌ی یک باشد و از آنجایی که $C_3, C_4 \cap \{x, y, z\} \neq \emptyset$ لذا ستون‌های C_3 و C_4 در سطر پنجم دارای درایه‌ی یک نیستند. فرض می‌کنیم ستون C_3 در سطر ششم دارای درایه‌ی یک باشد و چون $C_2, C_4 \cap \{x, y, z\} \neq \emptyset$ لذا ستون‌های C_2 و C_4 در این سطر دارای درایه‌ی یک نیستند و مشابه دلایل فوق ستون C_4 نیز در سطر هفتم دارای درایه‌ی یک است به‌طوری‌که ستون‌های دیگر در این سطر دارای درایه‌ی یک نیستند.

	C_1	C_2	C_3	C_4
r_1	۱	۱	۱	
r_2	۱	۱		۱
r_3	۱		۱	۱
r_4	۱			
r_5		۱		
r_6			۱	
r_7				۱

از آنجایی که ستون C_2 دارای وزن چهار است لذا باید درایه‌ی یک دیگری را به این ستون اضافه کرد. این درایه در سطر سوم نمی‌تواند قرار بگیرد. زیرا در این صورت $\lambda_{C_1, C_2} = 3$ می‌شود که این تناقض است. لذا این درایه باید در سطر هشتم قرار بگیرد. از طرفی برای کامل بودن $I(C_2)$ نیاز به دو ستون دیگری داریم که با C_2 در دو درایه اشتراک داشته باشد. حال اگر $q = 4$ باشد C_2 باید با C_3 و C_4 اشتراک داشته باشد و این فقط در سطر هشتم امکان‌پذیر است. در این صورت C_2, C_3 و C_4 کامل می‌شوند. مانند شکل زیر

	C_1	C_2	C_3	C_4
r_1	<u>۱</u>	۱	۱	
r_2	۱	۱		<u>۱</u>
r_3	۱		<u>۱</u>	۱
r_4	۱			
r_5		۱		
r_6			۱	
r_7				۱
r_8		<u>۱</u>	۱	۱

حال با حذف درایه‌های مشخص شده در شکل فوق مشاهده می‌شود که هر دو ستون حداکثر دارای یک اشتراک هستند بنابراین در M'_1 ، $\bar{\lambda} = 1$ است. حال اگر ستون‌های M'_1 را گسترش دهیم و مراحل فوق را انجام دهیم به ماتریس M_1 می‌رسیم که $\bar{\lambda} = 1$ است. از آنجایی که در M_2 ، $\bar{\lambda} = 1$ است و همچنین به‌ازای هر ستون C_1 از M_1 و ستون C_2 از M_2 داریم $\lambda_{C_1, C_2} \leq 1$. لذا ماتریس M به یک ماتریس ۲-مجزا با $\bar{\lambda} = 1$ و وزن سه یا چهار کاهش پیدا کرده است. لذا با استفاده از لم (۵.۳.۶)، $t - 1 - r = 8$ پس $n \leq \lfloor \frac{10 \times 8}{3 \times 2} \rfloor = 13$.

حال اگر $q \geq 5$ راه‌های دیگری برای کامل کردن $I(C_2)$ ، $I(C_3)$ و $I(C_4)$ وجود دارد. با استفاده از لم (۷.۴.۶) داریم $n'_1 \geq q \geq 5$. فرض می‌کنیم C_5 ستون جدیدی در M''_1 باشد. از آنجا که M''_1 مینیمال است لذا $I(C_5)$ باید با مجموعه‌ی $\{(1, 2), (1, 3), (2, 3)\}$ اشتراک داشته باشد. بدون اینکه به کلیت مسأله خللی وارد شود فرض می‌کنیم $(1, 2) \in I(C_5)$ از آنجا که ستون C_5 با ستون‌های C_1 ، C_2 ، C_3 و C_4 اشتراک دارد لذا بنابر لم (۸.۴.۶) نباید ستون C_5 با چهارمین درایه‌های یک این ستون‌ها اشتراک داشته باشد لذا دو درایه یک این ستون در سطرهاى جدید هشت و نه قرار می‌گیرند.

	C_1	C_2	C_3	C_4	C_5
r_1	۱	۱	۱		۱
r_2	۱	۱		۱	۱
r_3	۱		۱	۱	
r_4	۱				
r_5		۱			
r_6			۱		
r_7				۱	
r_8					۱
r_9					۱

با بررسی ارتباط ستون C_5 با بقیه ستون‌ها متوجه مقارن بودن دو درایه‌ی اول خواهیم شد. فرض می‌کنیم $(1, 8) \in I(C_5)$ ، که برای کامل بودن $I(C_5)$ نیاز به ستون C_6 با دو درایه‌ی یک در سطرهاى یک و هشت هستیم. از آنجا که ستون C_6 تنها با ستون C_4 اشتراک ندارد لذا ستون C_6 در سطر هفتم می‌تواند قرار داشته باشد و چهارمین درایه‌ی یک آن را در سطر جدید r_{10} قرار می‌دهیم و برای کامل بودن $I(C_3)$ ، $I(C_2)$

و $I(C_4)$ نیاز به حداقل یک سطر بیشتر داریم.

	C_1	C_2	C_3	C_4	C_5	C_6
r_1	۱	۱	۱		۱	۱
r_2	۱	۱		۱	۱	
r_3	۱		۱	۱		
r_4	۱					
r_5		۱				
r_6			۱			
r_7				۱		۱
r_8					۱	۱
r_9					۱	
r_{10}						۱

□

$$\text{لم } ۱۱.۴.۶. [۲۰] \quad n(3, 13) = 13$$

اثبات. ماتریس $M_{13 \times n}$ ، ۳-مجزا را در نظر می گیریم.

۱. اگر $\bar{\lambda} = 1$ بنا بر لم (۶.۳.۶) چون $d + 1 = 4$ لذا $d = 3$ و $N(d) = 16$ فرض می کنیم $n(3, 13) = k > 13$.

لذا یک ماتریس $M_{13 \times 14}^*$ ، ۳-مجزا داریم. حال با افزودن یک گروه تک عضوی به عنوان سطر چهاردهم

که در ستون پانزدهام دارای درایه‌ی یک است یک ماتریس $M_{14 \times 15}'$ ، ۳-مجزا داریم که بنا به لم (۱.۲.۶)،

$N(d) \leq 14$ که این تناقض با $N(d) = 16$ است. لذا $k \leq 13$ و از طرفی $n(3, 13) \geq 13$ لذا تساوی

برقرار است.

۲. حال فرض می کنیم دو ستون C و C' وجود دارند که $\lambda_{C_1, C_2} > 1$. در این صورت C ستون منفرد است یا

دارای وزن حداقل پنج است.

الف. اگر ستون C منفرد باشد با حذف آن و سطر متناظر با آن و از نتیجه (۶.۴.۶) داریم:

$$N(3) \geq 13 \Rightarrow n(3, 12) = 12,$$

و از لم (۳.۳.۶) داریم:

$$n \leq 1 + n(3, 12) = 1 + 12 = 13.$$

ب. حال اگر ستون C دارای وزن حداقل پنج باشد ستون C و سطرهاى برخوردکننده به آن را حذف می‌کنیم.

ماتریس حاصل $M'_{q \times (n-1)}$ ، ۲- مجزا با $q \leq 8$ است. چون $n(2, 8) = 8$ لذا $n - 1 \leq 8$ یعنی $n \leq 9$.

□

لم ۱۲.۴.۶. [۲۰] $n(3, 14) = 14$

اثبات. ماتریس $M_{14 \times n}$ ، ۳- مجزا را در نظر می‌گیریم.

۱. اگر $\bar{\lambda} = 1$ باشد بنابر لم (۶.۳.۶) چون $d + 1 = 4$ لذا $N(d) = 16$. فرض می‌کنیم $n(3, 14) = k$

که $k > 14$. لذا یک ماتریس $M_{14 \times 15}^*$ ، ۳- مجزا داریم. حال با افزودن یک گروه تک عضوی در سطر

پانزدهم که در ستون شانزدهم دارای درایه‌ی یک است یک ماتریس $M'_{15 \times 16}$ ، ۳- مجزا داریم، که بنا به لم

(۱.۲.۶)، $N(d) \leq 15$ که این تناقض با $N(d) = 16$ دارد لذا $k \leq 14$. از طرفی $n(3, 14) \geq 14$

پس $n(3, 14) = 14$.

۲. حال فرض می‌کنیم دو ستون C و C' وجود دارند که $\lambda_{C, C'} > 1$. لذا یا C منفرد است و یا دارای وزن

حداقل پنج است.

الف. اگر C ستون منفرد باشد با استفاده از لم (۳.۳.۶) و (۱۱.۴.۶)، $n \leq 1 + n(3, 13) = 14$ است.

ب. اگر C دارای وزن حداقل پنج باشد، ستون C و سطرهاى برخوردکننده به آن را حذف می‌کنیم لذا یک ماتریس

$M''_{q \times (n-1)}$ با $q \leq 9$ و ۲- مجزا داریم. با استفاده از لم (۳.۳.۶) و (۹.۴.۶) $n \leq 1 + n(2, 9) = 13$

است.

□

قضیه ۱۳.۴.۶. $[۲۰]$ $N(۳) = ۱۶$.اثبات. از آنجایی که $d + ۱ = ۲$ ، لذا با استفاده از قضیه (۴.۲.۶)، $N(۳) \leq ۱۶$. بنابراین کافی است تاثابت کنیم $N(۳) \geq ۱۶$ یا $n(۳, ۱۵) = ۱۵$. ماتریس ۳-مجزا، $M_{۱۵ \times n}$ را در نظر می‌گیریم با استفاده از قضیه(۶.۳.۶) فرض می‌کنیم دو ستون C و C' وجود دارند که $\lambda_{C,C'} > ۱$.الف. اگر C منفرد باشد با استفاده از لم (۳.۳.۶) و (۱۲.۴.۶) داریم:

$$n \leq ۱ + n(۳, ۱۴) = ۱۵.$$

ب. اگر C دارای وزن حداقل چهار باشد با استفاده از لم (۱.۴.۶) و (۱۰.۴.۶) داریم:

$$n \leq ۱ + n(۲, ۱۰) \leq ۱ + ۱۳ = ۱۴.$$

□

به راحتی می‌توان نشان داد $N(۴) = ۲۵$.

مراجع

- [1] H. B. Chen, D. Z. Du, and F. K. Hwang. Learning a hidden subgraph. *SIAM J. Discrete Math*, 18:697–712, 2005. [14](#)
- [2] H. B. Chen, D. Z. Du, and F. K. Hwang. An unexpected meeting of four seemingly unralated problems: graph testing ,DNA complex screening , superimposed codes and secure key distribution. *J Comb Opt* , to appear, 2007. [11](#)
- [3] H. D. Chen and F. K. Hwang. Exploring the missing link among d-separable,d-separable and d-disjunct matrices. *Applied Mathematics*, 133:662–664, 2007. [15](#)
- [4] Hong-Bin Chen, Yongxi Cheng, Qian He and Chongchong Zhong. Transforming an error-tolerant separable matrix to an error-tolerant disjunct matrix. *Discrete Applied Mathematics*, 157:387–390, 2009. , [20](#), [21](#), [22](#), [23](#), [24](#)
- [5] Hong-Bin Chen and Hung-Lin Fu. nonadaptive algorithms for threshold group testing. *Discrete Applied Mathematics*, 157:1581–1585, 2009. , [55](#), [57](#), [58](#), [59](#), [60](#)
- [6] Hong-Bin Chen and Frank K. Hwang. A survey on nonadaptive group testing algorithms through the angle of decoding. *J Comb Optim*, 15:49–59, 2008. , [45](#), [46](#), [48](#), [50](#), [51](#), [52](#)
- [7] P. Damaschke. Threshold group testing , general theory of information transfer and combinatorics . *LNCS*, 4123:707–718, 2006. [43](#), [58](#)
- [8] Annalisa De Bonis. New combinatorial structures with applications to efficient group testing with inhibitors. *J Comb Optim*, 15:77–94, 2008. , [63](#), [65](#), [66](#), [67](#), [68](#), [73](#), [75](#)
- [9] R. Dorfman. The detection of defective members of large populations. *Ann. Math. Statist*, 14:436–440, 1943. [1](#)
- [10] D. Z. Du and F. K. Hwang. Minimizing a combinatorial functin. *SIAM J. Algebraic Discrete Methods*, 3:523–528, 1982. [8](#)
- [11] D. Z. Du and F. K. Hwang. Combinatorial group testing and its aplicatios. *World Scientific*, 2000. [13](#), [30](#)

- [12] D. Z. Du and F. K. Hwang. *pooling Designs and nonadaptive group testing important tools for DNA sequencing*. World scientific. 2006. [10](#), [15](#), [20](#), [98](#)
- [13] Ding Zhu Du and Frank K. Hwang. *Combinatorial group testing and it's applications*. Series on Applied Mathematics. Co. Pte. Ltd., 2nd edition, 2000. [1](#), [8](#)
- [14] A. G. Dyachkov and V. V. Rykov. A survey of superimposed codes theory. *Problems Control Inform Theory*, 12:229–242, 1983. [8](#), [89](#)
- [15] A. G. Dyachkov, A. J. Macula, and V. V. Rykov. New applications and results of superimposed code theory arising from the potentialities of molecular biology. *Numbers and combinatorics , Proc., " Numbers, information , and complexity:, Conference,Bielefeld Germany ,Ingo Althofer , et al.eds.,Kluwer academic publishers*, pages 265–282, 2000. [62](#)
- [16] A. G. Dyachkov, A. J. Macula, and P. A. Vilenkin. Nonadaptive and trivial two-stage group testing with error-correcting $(d,e+1)$ -disjunct inclusion matrices. *Entropy,Search,Complexity*, pages 71–83. , [17](#), [75](#), [76](#), [77](#), [78](#), [79](#), [80](#)
- [17] P. Erdős and L. Moser. *Problem 35,Proceeding on the conference of combinatorial structures and their application*. Calgary. Gordon and breach, new yourk, 1970. [15](#)
- [18] W. Feller. *An Introduction to Probability Theory and it's Applicatins*. Series on Applied Mathematics. 1950. [2](#)
- [19] M. C. Hu, F. K. Hwang, and J. K. Wang. A boundary problem for group testing. *SIAM J. Algebraic Discrete Methods*, 2:81–87, 1981. [8](#)
- [20] S. H. Huang and F. K. Hwang. When is individual testing optimal for nonadaptive group testing?. *Siam J. Discrete Math*, 14:540–548, 2001. , [82](#), [83](#), [85](#), [86](#), [87](#), [88](#), [89](#), [90](#), [92](#), [93](#), [96](#), [97](#), [98](#), [99](#), [103](#), [104](#), [105](#)
- [21] T. Huang and C. Wang. Pooling spaces and non-adaptive pooling designs. *Discrete Math*, 282:163–169, 2004.
- [22] F. K. Hwang. A minimax procedure on group testing problems. *Tamkang J. Math.*, 2:39–44, 1971. [8](#)
- [23] F. K. Hwang and Y. C. Liu. Error-tolerant pooling designs with inhibitor. *J Comput Biol*, 10:231–236, 2003. [11](#), [47](#)
- [24] W. H. Kautz and R. R. Singleton. Nonrandom binary superimposed codes. *IEEE Trans.inform.Theory*, 10:363–377, 1964. [7](#), [15](#)
- [25] C. H. Lie. a sequential method for screening exprimental variables. *J.Amer.Statist.Assoc*, 57:455–477, 1962. [3](#)

- [26] C. C. Lindner and C. A. Rodger. *Design Theory*. CRC Press. New York, 1997. [83](#)
- [27] L. Lovász. On the ratio of optimal integral and fractional covers. *Discret math*, 13:383–390, 1975. [68](#)
- [28] A. Macula. Non-adaptive group testing with error-resistant d-disjunct matrices. *Discrete Applied Mathematics*, 80:217–282, 1998. [62](#)
- [29] A. Macula and P. A. Vilenkin. Constructions of generalized superimposed codes based on incident structures ,with vilenkin, p. a.. *2000 IEEE Proceedings of International symposium on Information Theory*, page 3, 200. [62](#)
- [30] Q. A. Nguyen and T. Zeisel. Bounds on constany weight binary superimposed codes. *problems control inform.Theory*, 17:223–230, 1988. [82](#)
- [31] P.S.S.N.V.P. Rao and Bikas K. Sinha. Improved bounds for group testing designs. *Journal of Statistical Planning and Inference*, 136:260–269, 2006. , [27](#), [28](#), [29](#), [30](#), [40](#), [41](#)
- [32] M. Sobel and P. A. Groll. Group testing to eliminate afficiently all defectives in a binomial sample. *Bell System Tech.J*, 28:1179–1252, 1959. [2](#)
- [33] J. Spencer. Minimal completely separating systems. *J. Combine.Theory*, 8:446–447, 1970. [8](#), [89](#)
- [34] D. C. Torney. Sets pooling designs. *Ann Comb*, 3:95–101, 1999. [11](#)
- [35] F. Vakil. Non-adaptive group testing procedures for hypergeometric problem. *Ph.D.Dissertatien , Temple University*, 1990. [33](#)
- [36] F. Vakil, M. Parnes, and D. Raghavarao. Group testing with at most two defectives when every item is included in exactly two group test. *Utilitas Math*, 38:161–164, 1990. [33](#)
- [37] R. Wilson. A diagonal form for the incident matrices of t-subsets vs.k-subsets. *European J. Conbin*, 6:609–615, 1990. [77](#)
- [38] A. Yakir. Inclusion matrix of k vs. l affine subspaces and a permutation module of the general affine group. *J. Combin. Theory Ser.A*, 63,no, 2:301–317, 1993. [62](#)
- [39] N. Alon, R. Beigel, S. Kasif, and B. Sudakov. Learning a hidden matching. *SIAM j Comput*, 33:487–501, 2004. [11](#)
- [40] W. J. Bruno, D. J. Balding, E. Knill, D. Bruce, C. Whittaker, N. Dogget, R. stalling, and D. C. Torney. Design of efficient pooling expriment. *Genomics*, 26:21–30, 1995. [10](#)

- [41] K. A. Bush, W. T. Federer, H. Pesotan, and D. Raghavarao. New combinatorial designs and their applications to group testing. *Ars combin*, 1980. [26](#)
- [42] F. H. Chang, H. L. Chang, and Hwang F. K. pooling designs for clon library screening in the inhibitor complex model. *to appear*, 2007. [11](#), [49](#)
- [43] H. B. Chen, H. L. Fu, and F. K. Hwang. An upper bound of the number of test in pooling designs for the error-tolerant complex model. *Opt Lett*, *to appear*, 2007. [11](#), [13](#)
- [44] A. De Bonis, L. Gasieniec, and U. Vaccaro. Optimal two - stage algorithms group testing problems. *SIAM Comput*, 34:1253–1270, 2005. [11](#)
- [45] A. De Bonis and U. Vaccaro. Improved algorithms for group testing with inhibitors. *Inform Process Lett*, 67:57–64, 1998. [11](#)
- [46] A. De Bonis and U. Vaccaro. constructions of generalized superimposed codes with applications to group testing and conflict resolution in multiple access channels. *Theor Comput Sci*, 306:223–243, 2003. [49](#)
- [47] A. G. D'yachkov, A. J. Macula, D. C. Torney, and P. A. Vilenkin. Two models of nonadaptive group testing for designing screening experiment. *Attkinson AC ,Muller WG (eds) Proceeding of the 6th international workshop in model oriented design and analysis*, pages 63–75, 2001. [47](#)
- [48] A. G. D'yachkov and V. V. Rykov. A survey of superimposed code theory. *Problems Control Inform.theory*12, 4:1–13, 1983. [64](#), [86](#)
- [49] A. G. D'yachkov, P. A. Vilenkin, A. J. Macula, and D. C. Torney. Families of finite sets in which no intersectio of l sets is covered by the union of s others. *J Comb Theory Ser A*, 99:195–218, 2001. [11](#)
- [50] P. Erdős, P. Frankl, and D. Füredi. Families of finite sets in which no set is covered by the union of r others. *Israel J. Math*, 51:79–89, 1985. [9](#)
- [51] M. Farach, S. Kannan, E. knill, and S. muthukrishnan. Group testing problem with sequences in experimental molecular biology. *Proceedings of the compression and complexity of sequences*, pages 357–367, 1997. [11](#)
- [52] A. J. Macula and L. J. Popyack. A group testing method for finding patterns in data. *Discret Appl Math*, 144: 149–157. [11](#)
- [53] C. J. Mitchell and F. C. Piper. Key storage in secure networks. 21:215–228, 1988. [12](#)

- [54] P. A. Pevzner and R. Lipshutz. towards DNA sequencing chips. *Mathematical Foundations of Computer Science*, 841:143–158, 1994. [10](#)
- [55] G. M. Saha and B. K. Sinha. some combinatorial aspect of designs useful in group testing experiments. *Unpublished manuscript*, 1980. [26](#)
- [56] G. M. Saha and B. K. Sinha. Some combinatorial aspect of designs useful in group testing experiments - an addendum. *Tach*, 1981. [26](#)
- [57] D. R. Stinson and R. Wei. Generalized cover-free families. 279:463–477, 2004. [50](#)
- [58] D. R. Stinson and R. Wei. Generalized cover-free families. *Discret Math*, 279:463–477, 2004. [13](#)
- [59] D. R. Stinson, R. Wei, and L. Zhu. Some new bounds for cover-free families. *J Comb Theory Ser A*, 90:224–234, 2000. [13](#)
- [60] C. A. Weideman and D. Raghavarao. Nonadaptive hypergeometric group testing designs for identifying at most two defectives. *Commun.Statist. Theory Methods*, pages 2997–3006, 1987. [85](#)
- [61] Hong-Bin Chen and Frank K. Hwang. Exploring the missing link among d-separable, $\bar{d}$ -separable and d-disjunct matrices. *Discrete Applied Mathematics*, 155:662–664, 2007. , [25](#)

فهرست الفبایی

- آزمون گروهی ، ۱
 آزمون گروهی آستانه ، ۴۴
 آزمون گروهی پویا ، ۵
 آزمون گروهی چند مرحله‌ای ، ۵
 آزمون گروهی ایستا ، ۵
 الگوریتم پویا ، ۴
 الگوریتم ایستا ، ۴
 اختلاف ، ۵۶
 بردار اشخاص بیمار ، ۶
 بردار نتیجه ، ۶
 بیمار ، ۲
 پوچگر ، ۱۱
 توزیع کلید امن ، ۱۳
 جستجوی گراف ، ۵۵
 خانواده پوشش آزاد ، ۱۳
 خطا آزاد ، ۱۳
 خنثی شدن ، ۱۰
 سیستم سه‌تایی اشتاینر ، ۹۸
 ضریب گاوسی ، ۷۸
 فاصله همینگ ، ۱۶
 کدهای ابر پوشش دهنده ، ۱۳
 گراف پنهان ، ۵۶
 گراف پترسن ، ۳۱
 $n(d, t)$ ، ۸
 $N(d)$ ، ۸
 $t(d, n)$ ، ۸
 $I(C)$ ، ۹۴
 $U(S)$ ، ۵
 $\|H\|$ ، ۵۶
 $N(d, q, n)$ ، ۶۵
 $[n]$ ، ۷۶
 $\binom{[n]}{k}$ ، ۷۶
 $\binom{[n]}{k}[q]^k$ ، ۷۷
 $\tau(H)$ ، ۶۹
 $H(X, Y)$ ، ۱۶
 $t(n, d, r, z)$ ، ۱۴
 $\delta(n, d, k, \Gamma, r)$ ، ۷۶
 $\pi(n, d, k, \Gamma, r)$ ، ۷۷
 $\gamma(n, d, k, \Gamma, r)$ ، ۷۸
 d - مجزا ، ۷
 d - جداشدنی ، ۷
 $\bar{d}$ - جداشدنی ، ۷
 (d, r) - مجزا ، ۱۲
 (d, z) - مجزا ، ۱۳
 (d, r, z) - مجزا ، ۱۴
 $(H_{\bar{r}} : d)$ - مجزا ، ۱۲
 (d, z) - جداشدنی ، ۱۶
 $(\bar{d}, z)$ - جداشدنی ، ۱۶
 d - قدرت تشخیص ، ۲۶
 u - کامل ، ۵۶
 u - ابرگراف ، ۵۶
 (k, m, w, n) - انتخابگر ، ۶۸
 (d, r) - ساختار پوچگر ، ۶۴
 $(0, 1)$ - ماتریس ، ۵
 آزمایش کلونی ، ۱۰

ماتریس شمول ، ۷۶

مثبت ، ۲

منفرد ، ۸۷

منفی ، ۲

وزن ، ۵

هم تافت ، ۱۱

واژه‌نامه فارسی به انگلیسی

d -disjunct	d - مجزا
d -separable	d - جداشدنی
$\bar{d}$ -separable	$\bar{d}$ - جداشدنی
$(d, r]$ -disjunct	$(d, r]$ - مجزا
(d, r, z) -disjunct	(d, r, z) - جداشدنی
$(H_{\bar{r}} : d)$ -disjunct	$(H_{\bar{r}} : d)$ - مجزا
(d, z) -separable	(d, z) - جداشدنی
$(\bar{d}, z)$ -separable	$(\bar{d}, z)$ - جداشدنی
d -detectpower	d - قدرت تشخیص
u -complet	u - کامل
u -hypergraph	u - ابرگراف
inhibitory (d, r) -design	(d, r) - ساختار پوچگر
(k, m, w, n) -selector	(k, m, w, n) - انتخابگر
binary matrices	$(0, 1)$ - ماتریس
clon librarie	آزمایش کلونی
group testing	آزمون گروهی
threshold group testing	آزمون گروهی آستانه
adaptive group testing	آزمون گروهی پویا
multi-stage group testing	آزمون گروهی چند مرحله‌ای
nonadaptive group testing	آزمون گروهی ایستا
sequential algorithm	الگوریتم ترتیبی
nonadaptive algorithm	الگوریتم ایستا
gap	اختلاف
vector of defective items	بردار اشخاص بیمار
result vector	بردار نتیجه
defective	بیمار
inhibitor	پوچگر
secure key distribution	توزیع کلید امن
graph search	جستجوی گراف
cover free familiy	خانواده پوشش آزاد
error-tolerant	خطا آزاد
hibridization	خنثی شدن

Steiner triple system	سیستم سه‌تایی اشتاینر
Gaussian coefficients	ضریب گاوسی
Hamming distance	فاصله همینگ
superimposed codes	کدهای ابر پوشش دهنده
Petersen graph	گراف پترسن
hidden graph	گراف پنهان
inclusion matrices	ماتریس شمول
positive	مثبت
isolated	منفرد
negative	منفی
weight	وزن
complex	هم‌تافت

Abstract

In this thesis, we present nonadaptive group testing. Also, three types of binary matrices have been found to be major tools in understanding and constructing pooling designs; we give their definitions. We shall show that nonadaptive group testing needs at least t group tests for identifying all the defective items from a group of t items having at most defective items. We use of graphs on group testing designs. Also we shall show that under restrictions on group size, optimal nonadaptive group testing can be constructed using Generalized Peterson Graphs. We present nonadaptive group testing algorithms, then those Generalize to error-tolerant model, presence of inhibitors, complex model and threshold model. We define new combinatorial structures with applications to efficient group testing with inhibitors. Let $N(d)$ denote the largest n for fixed d for which individual testing is optimal. In the rest of this thesis, we answer to query as follow; when is individual testing optimal for nonadaptive group testing? We shall show that $N(d) = (d + 1)$ for $d = 1, 2, 3, \dots$.

Keywords: *nonadaptive group testing, disjunct matrix.*



Shahrood University
Faculty of Mathematical Sciences
Department of Mathematics

M.S Thesis

Group testing

By:
Somayeh Khajeh

Supervisor:
Dr. Nader Jafari Rad
Dr. Meysam Alishahi

September 2012