دانشگاه صنعتی اصفهان

دانشکده علوم ریاضی

طرح‌های بلوکی بهینه در مدل‌های با متغیر کمکی

پایان نامه کارشناسی ارشد (آمار ریاضی)

میثم نجفی جعفرآبادی

استاد راهنما

دکتر سعید پولادساز

۱۳۸۹

دانشگاه صنعتی اصفهان

دانشکده علوم ریاضی

پایان نامه کارشناسی ارشد (آمار ریاضی) آقای میثم نجفی جعفرآبادی

تحت عنوان

طرح‌های بلوکی بهینه در مدل‌های با متغیر کمکی

در تاریخ ۸۹/۱۱/۲۵ توسط کمیته تخصصی زیر مورد بررسی و تصویب نهائی قرار گرفت.

۱– استاد راهنمای پایان‌نامه	دکتر سعید پولادساز
۲– استاد مشاور پایان‌نامه	دکتر ایرج کاظمی
۳– استاد داور ۱	دکتر هوشنگ طالبی (دانشگاه اصفهان)
۴– استاد داور ۲	دکتر محمد بهرامی

سرپرست تحصیلات تکمیلی دانشکده دکتر اعظم اعتماد

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و نوآوری‌های ناشی از
تحقیق موضوع این پایان‌نامه متعلق به دانشگاه صنعتی اصفهان است.

فهرست مطالب

فصل اول مقدمه ۱

۱–۱ طرح آزمایش چیست ؟ ۱

۱–۲ اصول پایه‌ای ۳

۱–۳ طراحی یک آزمایش ۴

۱–۴ آشنایی با طرح کاملاً تصادفی شده ۵

۱–۵ طرح‌های بهینه با متغیرهای کمکی ۵

۱–۶ تاریخچه بررسی ساختار طرح‌های بهینه با متغیرهای کمکی ۷

فصل دوم آشنایی با طرح‌های بلوکی و بعضی روش‌های ساختن آن‌ها ۹

۲–۱ طرح‌های بلوکی کاملاً تصادفی شده ۹

۲–۲ طرح‌های بلوکی ناقص ۱۰

۲–۲–۱ نماد و مدل ۱۱

۲–۲–۲ معادلات نرمال و نرمال کاهش یافته ۱۱

۲–۳ طرح های بلوکی ناقص متعادل ۱۵

۲–۴ روش‌های تفاضلی در ساختن طرح‌های بلوکی ناقص متعادل $(BIBD)$ ۱۶

۲–۴–۱ توسعه‌ی دوره‌ای مجموعه‌های تفاضلی ۱۶

۲–۴–۲ روش تفاضل‌های پی‌درپی متقارن ۲۰

۲–۴–۳ فرمول بندی بوسیله‌ی نظریه میدان گالوا ۲۷

۲–۵ طرح‌های بلوکی ناقص جزئی متعادل $(PBIBD)$ ۲۸

۲–۵–۱ نوع وابستگی ۲۸

۲–۵–۲ ماتریس‌های وابستگی ۲۹

۲–۵–۳ حل معادلات نرمال کاهش‌یافته ۳۰
۲–۵–۴ پارامترهای نوع دوم ۳۱
۲–۶ دسته‌بندی طرح‌های بلوکی ناقص جزئی– متعادل با m رده‌ی وابستگی ۳۵
۲–۶–۱ طرح‌های $PBIBD(۲)$ گروهی تقسیم‌پذیر ۳۶
۲–۶–۲ طرح‌های $PBIBD(۲)$ مثلثی ۳۸
۲–۶–۳ طرح‌های $PBIBD(۲)$ از نوع مربع لاتین ۳۹
۲–۷ آشنایی با بعضی روش‌های ساختن طرح‌های $PBIBD$ ۴۱
۲–۷–۱ ساختن طرح‌های $PBIBD(۲)$ گروهی تقسیم‌پذیر ۴۱
۲–۷–۲ ساختن طرح‌های $PBIBD(۲)$ مثلثی ۴۷
۲–۷–۳ ساختن طرح‌های $PBIBD(۲)$ نوع مربع لاتین ۴۸

فصل سوم طرح‌های آزمایشی بهینه برای مدل‌های همراه با متغیرهای کمکی ۵۰
۳–۱ مقدمه ۵۰
۳–۲ طرح‌های رگرسیون بهینه با عرض از مبدأهای یک طرفه دسته بندی شده ۵۲
۳–۳ طرح‌های رگرسیون بهینه با عرض از مبدأهای دو طرفه دسته بندی شده ۵۶
۳–۳–۱ طرح‌های رگرسیون بهینه در ساختار یک طرح بلوکی تصادفی ۵۷
۳–۴ رگرسیون بهینه با متغیرهای کمکی در ساختار یک طرح $BIBD$ ۶۵

فصل چهارم طرح‌های آزمایشی بهینه با متغیرهای کمکی در ساختار یک طرح بلوکی ناقص متعادل ۷۰
۴–۱ مقدمه ۷۰
۴–۲ طرح‌های $BIBD$ ساخته شده به‌وسیله‌ی روش تفاضلی بوس ۷۲
۴–۳ طرح‌های $BIBD$ دلخواه ۸۷

فصل پنجم طرح‌های آزمایشی بهینه با متغیرهای کمکی در ساختار یک طرح بلوکی ناقص جزئی–متعادل ۸۹
۵–۱ مقدمه ۸۹
۵–۲ طرح‌های بهینه با متغیرهای کمکی در ساختار طرح گروهی تقسیم پذیر منفرد . . . ۹۱
۵–۳ طرح‌های بهینه با متغیرهای کمکی در ساختار طرح گروهی تقسیم پذیر نیمه‌منظم . ۹۵
۵–۴ طرح‌های بهینه با متغیرهای کمکی در ساختار طرح گروهی تقسیم پذیر منظم . . . ۹۸

۵–۵ طرح‌های بهینه با متغیرهای کمکی در ساختار طرح $L_i - PBIBD$ ۱۰۲
۵–۶ طرح‌های بهینه با متغیرهای کمکی در ساختار طرح‌های مثلثی ۱۱۱

پیوست الف (آشنایی با ماتریس هادامارد) ۱۱۸

پیوست ب (آشنایی با مربع‌های لاتین دوبه‌دو متعامد و آرایه‌های متعامد) ۱۲۱

پیوست ج (آشنایی با میدان‌های متناهی و میدان‌های گالوا) ۱۲۴

پیوست د (جبر ماتریس‌ها) ۱۳۰

فهرست اسامی ۱۳۳

واژه‌نامه فارسی به انگلیسی ۱۳۵

واژه‌نامه انگلیسی به فارسی ۱۳۹

مراجع ۱۴۳

چکیده:

در طرح آزمایش‌ها گاهی با مسائلی مواجه می‌شویم که متغیر پاسخ با یک یا چند متغیر رابطه‌ی خطی دارد که به این متغیرها، متغیرهای کمکی می‌گویند. براساس شرایط آزمایش این متغیرهای کمکی به ۲ دسته زیر تقسیم می‌شوند.
(۱) حالتی که این متغیرهای کمکی، متغیرهای تصادفی می‌باشند و تحت کنترل آزمایشگر نیستند ولی می‌توان آن‌ها را همراه با متغیر پاسخ مشاهده کرد.
(۲) حالتی که این متغیرهای کمکی، متغیرهای غیر تصادفی می‌باشند و تحت کنترل آزمایشگر هستند که مقادیر آن‌ها را آزمایشگر مشخص می‌کند.
هنگامی که متغیر پاسخ با یک یا چند متغیر کمکی رابطه‌ی خطی دارد، استفاده از طرح‌های با متغیر کمکی مفید واقع می‌شود. با استفاده از این طرح‌ها میانگین مربعات خطا کاهش می‌یابد و باعث آشکار شدن اختلاف‌های واقعی حاصل از تیمارها در متغیر پاسخ می‌شود. در این پایان نامه متغیرهای کمکی را از نوع دوم که غیر تصادفی و قابل کنترل می‌باشند در نظر می‌گیریم و به دنبال یافتن طرح‌های بهینه با متغیر کمکی در ساختار طرح‌های بلوکی می‌باشیم به‌طوری‌که کاراترین برآورد را برای پارامترهای رگرسیونی مدل داشته باشیم.
در فصل اول به معرفی ابتدایی طرح‌های با متغیر کمکی می‌پردازیم و همچنین تاریخچه‌ای از توجه به این نوع طرح‌ها بیان می‌شود. در فصل دوم انواع طرح‌های بلوکی را معرفی و به بررسی ساختاری این طرح‌ها می‌پردازیم و بعضی روش‌های ساختن دسته‌ای از این طرح‌ها را بیان می‌کنیم.
در فصل های سوم، چهارم و پنجم به بیان قضیه‌هایی برای انواع طرح‌های بلوکی، در مورد ساختن ماتریس مقادیر متغیرهای کمکی موجود در مدل تحت شرایطی می‌پردازیم که کاراترین برآورد را برای پارامترهای رگرسیونی در ساختار طرح‌های بلوکی داشته باشیم.

کد رده‌بندی : اولیه ۶۲K۰۵، ثانویه ۶۲K۱۰

کلمات کلیدی : طرح‌های بلوکی، متغیرهای کمکی، طرح‌های بهینه، آرایه متعامد، ماتریس هادامارد، حاصل ضرب کاتری—رائو، حاصل ضرب کرانکر

فصل ۱

مقدمه

۱–۱ طرح آزمایش چیست ؟

طرح آزمایش را می‌توان با نقشه‌های ساختمانی که یک مهندس ساختمان برای احداث یک خانه ارائه می‌دهد مقایسه کرد. با وجودی که مهندس می‌تواند با به کارگیری خلاقیت خود نقشه‌های متنوعی ارائه دهد، اما موظف است که نیازهای اساسی ساکنین آینده‌ی ساختمان یا کارفرما را برآورده سازد. برای این منظور چندین طرح مختلف ارائه داده و از بین آن‌ها با توجه به تمام جوانب امر طرح مورد نظر انتخاب می‌شود. در طرح ریزی یک آزمایش، طراح در نقش مهندس و آزمایشگر در نقش صاحب آینده ساختمان یا کارفرما است، و در واقع آزمایشگر است که تصمیم نهایی را درباره‌ی آزمایش اتخاذ می‌کند.

بدون وجود یک طرح آزمایش مناسب نمی‌توان فرض‌های بالقوه سودمند را با درجه‌ی قابل قبولی از دقت آزمود. قبل از رد کردن فرضی مربوط به زمینه‌ای تحقیقاتی، باید ساختار آزمایش را مورد بررسی قرار داد و اطمینان حاصل کرد که آیا آزمایش، یک آزمون واقعی از فرض را فراهم ساخته است ؟

به معنای واقعی کلمه، آزمایش یک آزمون است. آزمایش طرح شده، یک آزمون یا دنباله‌ای از آزمون‌ها است که در آن‌ها تغییرات مورد نظر در متغیرهای ورودی فرایند یا سیستم به قسمی اعمال می‌شوند که بتوانیم علل تغییرات در پاسخ خروجی را مشاهده و مشخص کنیم.

اهداف آزمون می‌تواند شامل موارد زیر باشد.

۱) تعیین متغیرهایی که بیشترین تأثیر را در پاسخ نتیجه‌ی آزمایش دارند.

۲) تعیین موقعیت متغیرهای مؤثر به طوری که تقریباً پاسخ همیشه نزدیک مقدار مطلوب اسمی باقی بماند.
۳) تعیین موقعیت متغیرهای مؤثر به طوری که اثرهای کنترل ناپذیر کمینه شوند.

به عنوان مثال، فرض کنید پژوهشگری به مطالعه‌ی اثر چند عامل مختلف مانند نور، دما، رطوبت و خاک در میزان رشد یک گیاه آپارتمانی علاقه‌مند باشد. در اینجا هدف پژوهشگر تعیین عوامل و سطوحی از آن عوامل است که برای این گیاه خاص بیشترین میزان باروری را ایجاد می‌کنند. پژوهشگر تصمیم می‌گیرد که تعدادی از نمونه‌های این گیاه را در سطوح مختلف این عوامل پرورش داده و میزان رشد آن‌ها را اندازه بگیرد. سپس با استفاده از متوسط میزان رشد مشاهده شده در نمونه‌ها در اثر هریک از سطوح عوامل مختلف، معین کند که کدام یک از عوامل و در چه سطحی بهترین می‌باشند.

قبل از انجام آزمایش باید توجه داشت، نتایجی که می‌توان استخراج کرد در سطح وسیع وابسته به روشی است که داده‌ها جمع آوری شده‌اند. برای روشن شدن موضوع، فرض کنید که پژوهشگر آزمایش بالا، از نمونه‌هایی با یک ویژگی، مثلاً از نهال‌هایی جوانتر برای پرورش در دماهای مختلف و از نهال‌هایی قدری مسن‌تر برای پرورش در رطوبت‌های مختلف استفاده کند. حال وقتی میانگین رشد آن‌ها را مقایسه می‌کند، قادر نیست بفهمد که چقدر از اختلاف مشاهده شده معلول عامل دما و رطوبت و چقدر معلول اختلاف سن یا جنس نهال‌ها است. پس روش جمع آوری داده‌ها می‌تواند بر نتایجی که از آزمایش استخراج می‌شوند، اثر کند.

روش‌های طرح آزمایش‌ها در بسیاری از زمینه‌ها کاربرد وسیع دارند. در واقع عمل آزمایش را می‌توان به صورت بخشی از فرایند علمی و به صورت یکی از راه‌های کسب اطلاعات از واقعیت‌ها دانست.

در رشته‌های مهندسی، طرح آزمایش ابزاری فوق‌العاده مهم برای اصلاح عملکرد فرآیندهای تولید است. کاربرد تکنیک‌های اولیّه‌ی طرح آزمایش در توسعه‌ی فرآیند می‌تواند نتایج زیر را داشته باشد.

۱) فرآیند را اصلاح کند.
۲) تغییر پذیری را کاهش داده و آن را به نیازهای مورد نظر یا اسمی نزدیکتر کند.
۳) زمان تولید را کاهش دهد.
۴) هزینه را تقلیل دهد.

روش‌های طرح آزمایش در فرآیند تولید نیز نقش عمده دارند. استفاده از طرح آزمایش در این زمینه‌ها می‌تواند موجب تولید فرآورده‌های مرغوبتر شود. فرآورده‌هایی که کارایی بیشتری داشته و قابل اعتمادند و هزینه‌ی تولید پایین‌تر و زمان تولید کوتاهتر دارند.

در این قسمت لازم است که به معرفی پارامترهای یک طرح آزمایش بپردازیم. اولین واژه‌ای که در طرح آزمایش ممکن است مورد توجه قرار گیرد، واژه‌ی تیمار است، که به هر سطح از عوامل مورد آزمایش تیمار می‌گوییم. مثلاً اگر در مثال قبل هریک از عوامل دارای دو سطح باشند آنگاه ۸ تیمار وجود دارد ، زیرا در

مجموع ۴ عامل نور، دما، رطوبت و خاک را برای بررسی داریم و هر کدام دارای دو سطح است.
واژه‌ی بعدی که باید به آن توجه کرد واحد آزمایش است. به هر مکان یا چیزی که تیمارها بر روی آن اعمال می شوند واحد آزمایش گوییم. در مثال قبل هر نهال یک واحد آزمایش است.[۳۵]

۱ – ۲ اصول پایه‌ای

اگر بخواهیم نتایجی با معنا و معتبر از داده‌ها استخراج کنیم، استفاده از روش‌های آماری در طرح آزمایش ضروری است. وقتی مسأله متضمن داده‌هایی است که شامل خطاهای آزمایشی هستند در این صورت روش آماری تنها رهیافت عینی تحلیل است. پس هر پژوهش استقرایی دارای دو وجه است، یکی طرح آزمایش و انجام آن و دیگری تحلیل آماری داده‌ها. این دو موضوع در ارتباط نزدیک با یکدیگرند، زیرا روش تحلیل مستقیماً به استفاده طرح استفاده شده بستگی دارد.
سه اصل پایه‌ای طرح آزمایش عبارتند از:
۱) تکرار.
۲) تصادفی کردن.
۳) بلوک‌بندی.
منظور از تکرار، تهیه‌ی مشاهدات متعدد در امتحان واحدهای آزمایش است. در مثال قبل، تکرار عبارت از پرورش نمونه‌هایی از گیاه در هریک از سطوح عوامل مختلف (تیمارها) است. مثلاً اگر برای هریک از عوامل نور، دما، رطوبت و خاک دو سطح و برای هریک از ترکیب‌های تیماری پنج تکرار در نظر بگیریم، آنگاه باید برای ملاحظه‌ی اثر تمام ترکیب‌های تیماری در رشد گیاه جمعاً $۸۰ = ۵ \times ۲^۴$ مشاهده داشته باشیم.
تکرار دو خاصیت دارد. اول این‌که پژوهشگر را قادر می‌سازد که برآوردی برای خطای آزمایش به‌دست آورد. این برآورد خطا مبنای واحد اندازه‌گیری در تعیین معنی‌دار بودن اختلاف مشاهده شده در داده‌ها است. وقتی می گوییم اختلاف معنی‌دار است یعنی اختلاف از آن حدی که خطاهای نمونه‌گیری و خطاهای آزمایشی می‌توانند ایجاد کنند تجاوز کرده است و بنابراین نمی‌تواند معلول خطا باشد. دومین خاصیت تکرار آن است که مثلاً، اگر برای برآورد اثر سطحی از یک عامل، از میانگین نمونه استفاده کنیم، آن‌گاه تکرار موجب برآوردی دقیق‌تر برای آن اثر می‌شود.
در استفاده از روش‌های آماری در طرح آزمایش‌ها، تصادفی کردن مسئله‌ای بنیادی است. منظور از تصادفی کردن آن است که تخصیص تیمارها به واحدهای آزمایش و ترتیبی که مطابق آن آزمایش‌ها انجام می‌شوند به تصادف تعیین شوند. تصادفی کردن درست آزمایش، به خارج کردن متوسط اثرهای

عوامل خارجی که می‌توانند در نتیجه‌ی نهایی تأثیر گذار باشند کمک می‌کند. مثلاً، فرض کنید در آزمایش بالا نمونه‌ها از نظر سن با یکدیگر اختلاف جزئی داشته باشند و اثر عامل نور بتواند تحت تأثیر سن نمونه‌ها قرار بگیرد. اگر تمام نمونه‌هایی که در سطح اول عامل نور پرورش یافته‌اند نسبت به آن‌هایی که در سطح دوم عامل نور پرورش پیدا کرده‌اند، مسن‌تر باشند، آن‌گاه ممکن است به غلط سطحی از نور را نسبت به سطح دیگر برتر بدانیم. تخصیص تصادفی نمونه‌ها به سطوح مختلف این مشکل را برطرف می‌کند.
بلوک‌بندی تکنیکی است که برای افزایش دقت آزمایش از آن استفاده می‌کنیم. استفاده از آن در آزمایش‌ها موجب می‌شود که علیرغم محدودیت‌های واقعی در انتخاب تصادفی واحدهای آزمایشی بتوانیم خطای آزمایش را کاهش دهیم. اصطلاح بلوک اشاره به متجانس بودن نسبی واحدهای مورد آزمایش دارد و محدودیت در تصادفی کردن کامل را نشان می‌دهد. در بسیاری از مسائل آزمایشی لازم است آزمایش به گونه‌ای طرح شود که بتوان تغییرات حاصل از منابع خارجی را با روشی معین کنترل کرد. مایلیم خطای آزمایش را تا حد ممکن کم کنیم یعنی، مایلیم تغییر پذیری بین واحدهای آزمایش را از خطای آزمایش جدا کنیم. بلوک بندی تکنیکی است که چنین خواسته‌ای را تامین می‌کند. مثلاً در بسیاری از موارد عملاً در انجام آزمایش متوجه می‌شویم که نمی‌توانیم تیمارها را تحت شرایط کاملاً یکسان به واحدهای آزمایش نسبت دهیم. یا ممکن است متوجه شویم در انجام آزمایش لازم است از اشخاص مختلفی استفاده کنیم که به دلیل تجربه و تبحر خاص خود عملکردهای متفاوت دارند. یا در انجام آزمایش نیاز به مواد خام داشته باشیم و متوجه شویم مواد خامی که در اختیار داریم کیفیتی متفاوت دارند. بالاخره ممکن است ناگزیر باشیم از ماشین‌های مختلفی برای اندازه‌گیری متغیر پاسخ استفاده کنیم. اشخاص، ماشین‌ها، مواد خام، زمان انجام آزمایش و نظایر این‌ها خود منابع بالقوه‌ی تغییر پذیری هستند. با بلوک‌بندی می‌توانیم این منابع خارجی تغییر پذیری را کنترل کنیم.[۳۵]

۱–۳ طراحی یک آزمایش

طرح کردن یک آزمایش عبارت است از تدارک انجام یک آزمایش که برای افزایش صحت و دقت اطلاعات انجام می‌گیرد و معمولاً برای یک طراحی آزمایش مناسب قدم‌های زیر برداشته می‌شود.
۱) بیان اهداف آزمایش
به‌طور مثال اثر سنجی و مقایسه‌ی روش‌های مختلف آموزش در یادگیری یا اثر سنجی انواع بذرها در میزان محصول گندم و مقایسه‌ی بذرها برای انتخاب بهترین بذر یا تعیین اثر مواد اولیه در کیفیت یک کالا و مقایسه‌ی انواع آن‌ها

۲) انتخاب تیمارهای آزمایشی (عوامل آزمایش)
ممکن است تیمار آزمایشی از یک عامل و یا از ترکیب عوامل مختلف باشد.
۳) تعیین عوامل غیر آزمایشی
۴) تعیین واحد آزمایشی
واحد آزمایشی باید به‌گونه‌ای باشد که تکرارپذیر باشد.
۵) تعیین تعداد تکرار
٦) تعیین صفتی که معرف اثر تیمارها می‌باشد (محصول آزمایش)
۷) شناخت طرح‌های آزمایشی مختلف برای انتخاب مناسب‌ترین طرح بر مبنای اطلاعات فوق.
۸) جمع‌آوری اطلاعات اضافی در انجام آزمایش برای استفاده از آن‌ها برای کسب اطلاعات بیشتر.

۱–۴ آشنایی با طرح کاملاً تصادفی شده

می‌توان گفت طرح کاملاً تصادفی شده ابتدایی‌ترین طرح در میان انواع طرح‌های آزمایشی می‌باشد. در یک طرح کاملاً تصادفی شده، تیمارها (سطوح عوامل مورد نظر) را به طور کاملاً تصادفی به مجموعه‌ی واحدهای آزمایش نسبت می‌دهیم و به این دلیل آن را طرح کاملاً تصادفی می‌نامیم. این طرح زمانی کارا است که بین واحدهای آزمایشی هیچ‌گونه منبع تغییر شناخته شده‌ی دیگری به جز تیمارها وجود نداشته باشد. مثلاً در استفاده از این طرح برای مطالعه‌ی میزان محصول لازم است قطعات مختلف زمین از نظر رطوبت، حاصلخیزی خاک، نور و غیره همگن باشند. به این دلیل طرح کاملاً تصادفی‌شده نسبت به ترتیب فیزیکی واحدهای آزمایش انعطاف‌پذیر است. اگر منبع تغییر دیگری به جز تیمارها بین واحدهای آزمایش موجود باشد در این صورت برای آشکار کردن تفاوت واقعی تیمارها باید از طرح‌های دیگری مانند طرح‌های بلوکی برای سنجش تفاوت تیمارها استفاده کرد که در فصل دوم به بررسی کامل طرح‌های بلوکی می‌پردازیم.

۱–۵ طرح‌های بهینه با متغیرهای کمکی

دو روش کلی مستقیم و غیرمستقیم(آماری) برای کنترل تغییرات حاصل از خطای آزمایشی وجود دارند. کنترل مستقیم، روش‌هایی از قبیل گروه‌بندی واحدهای آزمایشی در طبقات همگن یا بلوک‌ها، افزایش یکنواختی شرایطی که آزمایش تحت آن‌ها انجام می‌شود و افزایش درستی اندازه‌گیری‌ها را دربرمی‌گیرد.

در ادامه به معرفی طرح‌هایی می‌پردازیم که در آن‌ها از کنترل غیر مستقیم یا آماری برای افزایش دقت آزمایش استفاده می‌شود.

در طرح آزمایش‌ها ممکن است متغیر پاسخ y با یک یا چند متغیر رابطه‌ی خطی داشته‌باشد، که این متغیرها را با z_i نشان می‌دهند و به آن‌ها متغیرهای کمکی می‌گویند.

براساس شرایط آزمایش این متغیرهای کمکی به ۲ دسته تقسیم می‌شوند:

۱) حالتی که متغیرهای کمکی z_i ها، متغیرهای تصادفی می‌باشند و تحت کنترل آزمایشگر نیستند ولی می‌توان آن‌ها را همراه y مشاهده کرد.

۲) حالتی که متغیرهای کمکی z_i ها، متغیرهای غیر تصادفی می‌باشند و تحت کنترل آزمایشگر هستند که مقادیر آن‌ها را آزمایشگر مشخص می‌کند.

در مواردی که یک یا چند متغیر کمکی با y رابطه‌ی خطی دارند، استفاده از طرح‌هایی با متغیر کمکی مفید واقع می‌شود. با استفاده از این طرح‌ها میانگین مربعات خطا کاهش می‌یابد و باعث آشکار شدن اختلاف‌های واقعی حاصل از تیمارها در متغیر پاسخ می‌شود.

در این پایان نامه متغیرهای کمکی را از نوع دوم که غیرتصادفی و کنترلی می‌باشند، در نظر می‌گیریم. در حالت بلوکی مدل طرح‌های با متغیر کمکی را به صورت زیر تعریف می‌کنیم:

$$(Y,\ \mu 1_{\mathrm{N}} + X_1\beta + X_2\tau + Z\gamma,\ \sigma^2 I)$$

که در مدل بالا، μ میانگین کل، σ^2 واریانس مشترک مشاهدات، $\beta_{b\times 1}$، $\tau_{v\times 1}$ و $\gamma_{c\times 1}$ به ترتیب نشان دهنده‌ی بردارهای اثرات بلوکی، اثرات تیماری و اثرات متغیرهای کمکی می‌باشند. $Y_{N\times 1}$ بردار مشاهدات ناهمبسته، X_1 و X_2 به ترتیب ماتریس‌های طرح متناظر با اثرات بلوکی و اثرات تیماری و Z ماتریس طرح متناظر با اثرات متغیرهای کمکی هستند. همچنین 1_N یک بردار از مرتبه N است، که همه عناصر آن برابر ۱ می‌باشد.

بدون از دست دادن کلیت، برای متغیرهای کمکی تبدیل (مکانی—مقیاسی) انجام می‌دهیم به طوری که داشته باشیم:

$$|z_{ij}| \leqslant 1$$

برای یافتن برآورد متعامد اثرات تیماری و متغیرهای کمکی، و همچنین برآورد متعامد اثرات بلوکی و اثرات متغیرهای کمکی شرط لازم و کافی آن است که:[۲۵]

$$Z'X_1 = \mathbf{0} \quad , \quad Z'X_2 = \mathbf{0} \qquad (1.1)$$

برای اینکه برآورد پارامترهای رگرسیونی کارا باشند، باید شرط زیر نیز برقرار باشد [۲۵] :

$$Z'Z = NI_c \qquad (۲.۱)$$

یک ماتریس متناظر با متغیرهای کمکی Z بهینه است اگر برای X_1 وX_2 داده شده شرایط (۱.۱) و (۲.۱) برقرار باشد.

به عبارت دیگر شرایط (۱.۱) و (۲.۱) را می‌توان به‌صورت زیر تعبیر کرد،

i) ستون‌های ماتریس Z باید بر ستون‌های ماتریس‌های X_1 وX_2 عمود باشند.

ii) ستون‌های ماتریس Z باید دوبه‌دو متعامد بوده و عناصر Z باید ± 1 باشند.

در این پایان نامه به دنبال مشخص کردن ماتریس مقادیر متغیرهای کمکی Z، در ساختار طرح‌های مختلف بلوکی می‌باشیم به‌طوری که در شرایط (i) و (ii) گفته شده برای ماتریس Z بهینه، صدق کند.

۱–۶ تاریخچه بررسی ساختار طرح‌های بهینه با متغیرهای کمکی

مسئله‌ی بررسی ساختار طرح‌های بهینه با متغیرهای کمکی را لوپز ترویا [۲۳] و [۲۴]، داس و همکاران [۱۴]، رائو و همکاران [۲۷]، داتا [۱۶]، داتا و همکاران [۱۷] و [۱۸] مورد توجه قرار دادند.

در سال ۱۹۸۲ لوپز ترویا [۲۳] و [۲۴] ابتدا مسئله انتخاب ماتریس مقادیر متغیرهای کمکی Z، در طرح کاملاً تصادفی شده (CRD) را بررسی کرد.

داس و همکاران [۱۴] در سال ۲۰۰۳ مسئله انتخاب ماتریس مقادیر متغیرهای کمکی Z، در ساختار طرح‌های بلوکی تصادفی (RBD) و مجموعه‌ای از طرح‌های بلوکی ناقص متعادل بهینه $(BIBD)$ را مورد مطالعه قرار دادند.

رائو و همکاران [۲۷] در سال ۲۰۰۳ مسئله ساختار طرح‌های CRD وRBD را دوباره بررسی کردند و راه‌حل‌هایی بااستفاده از آرایه متعامد آمیخته ارائه کردند که در نتیجه، این عمل موجب بهبود نگرش به مسئله انتخاب ماتریس مقادیر متغیرهای کمکی Z، شده است.

داس وهمکاران [۱۴] در سال ۲۰۰۳ طرح‌های بلوکی ناقص متعادل متقارن بهینه $(SBIBD)$ را با پارامترهای $(b = v، r = k، \lambda)$ که از طریق روش تفاضلی بوس ساخته می‌شوند و مجموعه‌ای از طرح‌های $BIBD$ را که با تکرار بلوک‌ها ساخته می‌شوند مورد مطالعه قرار دادند.

داتا [۱۶] در سال ۲۰۰۴ همچنین مسئله انتخاب ماتریس مقادیر متغیرهای کمکی Z، را در مجموعه‌ی دیگری از طرح‌های $BIBD$ که از طریق روش تفاضلی بوس ساخته می‌شوند و به همراه بعضی طرح‌های

$BIBD$ دلخواه مورد بررسی قرار داد. پس از آن داتا و همکاران [۱۷] در سال ۲۰۰۷ مجموعه‌ای از طرح‌های $SBIBD$ را که از طریق تصویر هندسی بدست می‌آمدند مورد بررسی قرار دادند.

مشخص کردن ساختار طرح بهینه با متغیر کمکی برای یک طرح بلوکی ناقص متعادل $(BIBD)$ دلخواه مشکل می‌باشد.

همان‌طور که ملاحظه می‌شود تا سال ۲۰۰۹ ساختار طرح‌های بلوکی ناقص جزئی–متعادل $(PBIBD)$ که کلاس بزرگی از طرح‌های IBD را شامل می‌شوند مورد توجه قرار نگرفته بودند، در حالی که این نوع از طرح‌ها کاربرد زیادی در عمل دارند.

در سال ۲۰۰۹ مجموعه‌ای از طرح‌های $PBIBD$ توسط داتا و همکاران [۱۸] مورد بررسی قرار گرفته‌اند که تنها از طریق روش تفاضلی بدست نمی‌آیند بلکه بوسیله روش‌های دیگری که توسط بوس و همکاران [٦] و زِلن [۳۴] و وارتِک [۳۳] گفته شده، بدست می‌آیند.

در این پایان نامه، در فصل دوم انواع طرح‌های بلوکی به همراه بعضی روش‌های ساختن این طرح‌ها را بیان می‌کنیم و خصوصاً طرح‌هایی را که در فصل‌های بعد مورد بررسی قرا می‌دهیم به طور کاملتری توصیف می‌کنیم.

در فصل سوم ساختار طرح‌های بلوکی بهینه بدون حضور متغیرهای کمکی را در نظر می‌گیریم و به مسئله‌ی انتخاب ماتریس مقادیر کمکی هنگامی که در این مدل‌ها متغیر کمکی وارد می‌شود می‌پردازیم. به گونه‌ای که در سرانجام یک طرح بلوکی بهینه با حضور متغیر کمکی داشته باشیم.

در فصل چهارم و پنجم به ترتیب ماتریس مقادیر کمکی را در دسته‌هایی از طرح‌های بلوکی ناقص متعادل و بلوکی ناقص جزئی–متعادل طوری مشخص می‌کنیم که یک طرح‌بهینه با حضور متغیر کمکی در ساختار این طرح‌ها وجود داشته باشد.

فصل ۲

آشنایی با طرح‌های بلوکی و بعضی روش‌های ساختن آن‌ها

در این فصل با انواع طرح‌های بلوکی شامل طرح‌های بلوکی کاملاً تصادفی شده ($CRBD$) و طرح‌های بلوکی ناقص (IBD) آشنا می‌شویم. در ادامه انواع طرح‌های بلوکی ناقص شامل طرح‌های بلوکی ناقص متعادل ($BIBD$) و طرح‌های بلوکی ناقص جزئی متعادل ($PBIBD$) را به همراه بعضی از روش‌های ساختن این طرح‌ها ارائه می‌دهیم.

۲–۱ طرح‌های بلوکی کاملاً تصادفی شده

در بسیاری از مسائل مربوط به آزمایش‌ها، لازم است آزمایش به گونه‌ای طرح شود که بتوان تغییرپذیری ناشی از منابع مزاحم را به صورتی سیستماتیک کنترل کرد. مثلاً، در آزمون سختی یک فلز، می‌خواهیم بدانیم که استفاده از چهار تیغه‌ی مختلف منجر به نتایج یکسان می‌شود یا خیر. ماشین سنجش با فشار آوردن تیغه‌ای به قطعات فلز عمل می‌کند، و از عمق حاصل سختی قطعه مشخص می‌شود. آزمایشگر تصمیم می‌گیرد که از هر تیغه چهار مشاهده داشته باشد. تنها یک عامل، نوع تیغه، وجود دارد و طرح تک عاملی کاملاً تصادفی‌شده عبارت است از تخصیص تصادفی هریک از $۱۶ = ۴ \times ۴$ اجرا به واحد آزمایش، که همان قطعه‌ی فلز تحت آزمون می‌باشد. در این حالت، استفاده از طرح کاملاً تصادفی‌شده

توصیه نمی‌شود. زیرا اگر قطعات فلز از نظر سختی اندکی تفاوت داشته باشند، که‌مثلاً ممکن است از اختلاف حرارت شمش‌های مختلف در زمان تولید ناشی شده باشد، آن‌گاه واحدهای آزمایشی (قطعات) می‌توانند در تغییرپذیری مشاهده شده در داده‌های سختی نقش داشته باشند. در نتیجه خطای آزمایش بازتاب خطای تصادفی و تغییرپذیری بین قطعات می‌باشد.

مایلیم خطای آزمایش را تا حد ممکن کاهش دهیم، یعنی مایل به تفکیک تغییرپذیری بین قطعات از خطای آزمایش هستیم. بدین منظور ۴ قطعه فلز اختیار می‌کنیم و هر تیغه را یک‌بار روی هریک از قطعات اعمال می‌کنیم. واژه‌ی «کامل» نشان می‌دهد که هر بلوک (قطعه) شامل تمام تیمارها (تیغه‌ها) است. با استفاده از این طرح، بلوک‌ها یا قطعات برای مقایسه‌ی تیغه‌ها، واحدهای آزمایشی همگن‌تری را بوجود می‌آورند. اگر واریانس خطای آزمایش در طرح کامل تصادفی‌شده و واریانس خطای آزمایش در طرح بلوک کامل تصادفی‌شده را به ترتیب با $\sigma^2_{e(CRD)}$ و $\sigma^2_{e(CRBD)}$ نشان دهیم انتظار داریم $\sigma^2_{e(CRBD)} < \sigma^2_{e(CRD)}$ باشد.[۱۵]

دریک طرح بلوکی کاملاً تصادفی شده با v تیمار در b بلوک، هر تیمار در داخل هریک از b بلوک یکبار رخ می‌دهد و بنابراین اندازه‌ی هریک از b بلوک برابر تعداد تیمارها می‌باشد و همچنین تعداد تکرار هریک از تیمارها برابر با تعداد بلوک‌ها می‌باشد.

۲–۲ طرح‌های بلوکی ناقص

ممکن است در بعضی آزمایش‌ها با استفاده از طرح‌های بلوکی کاملاً تصادفی شده قادر به اجرای کلیه‌ی ترکیب‌های تیماری در هر بلوک نباشیم. چنین وضعیت‌هایی معمولاً به دلیل کمبود وسایل آزمایشی یا اندازه‌ی فیزیکی بلوک پیش می‌آید یعنی با حالتی مواجه می‌شویم که تعداد تیمارها از اندازه‌ی بلوک‌ها بیشتر می‌باشد. در این نوع مسائل از طرح‌های بلوکی ناقص استفاده می‌کنیم. مثلاً فرض کنید در آزمایش آزمون سختی فلز، به دلیل‌محدود بودن اندازه‌ی قطعات آزمایشی، تنها آزمون سه تیغه روی هر قطعه‌ی آزمایشی مقدور باشد. بنابراین هر تیغه‌ای را نمی‌توان با هر قطعه‌ای امتحان کرد.

در ادامه مدل یک طرح بلوکی ناقص با v تیمار در b بلوک را در حالت کلی که تعداد تکرار تیمارها متفاوت و اندازه‌ی هریک از بلوک‌ها نیز متفاوت می‌باشد، بیان می‌کنیم.

۲–۲–۱ نماد و مدل

فرض کنید v تیمار با تکرارهای $r_1, r_2, \ldots, r_v$ و b بلوک با اندازه‌های $k_1, k_2, \ldots, k_b$ را در نظر بگیریم. یعنی در بلوک j تعداد k_j واحد آزمایش قرار دارد. لذا داریم :

$$\sum_{i=1}^{v} r_i = \sum_{j=1}^{b} k_j = n$$

به طوری که n تعداد کل مشاهدات می‌باشد.

اکنون برای یک طرح بلوکی ناقص، مدل زیر را برای پاسخ y_{ij} در نظر می‌گیریم:

$$y_{ijl} = \mu + \tau_i + \beta_j + e_{ijl} \qquad (۱.۲)$$

$$l = 1, 2, \ldots, n_{ij}, \quad i = 1, 2, \ldots, v, \quad j = 1, 2, \ldots, b$$

به‌طوری که τ_i نمایانگر اثر تیمار i ام و β_j نمایانگر اثر بلوک jام و e_{ijl} خطای مربوط به مشاهده‌ی y_{ijl} (مشاهده‌ی مربوط به تیمار i ام در بلوک j ام و تکرار l ام) می‌باشد و n_{ij} نمایانگر تعداد تکرارهای تیمارهای i ام در بلوک j ام می‌باشد.

فرض می‌کنیم e_{ijl} ها، متغیرهای تصادفی مستقل و هم‌توزیع با میانگین صفر و واریانس σ_e^2 باشند.

حال مدل(۱.۲) را می‌توان به صورت ماتریسی زیر نوشت:

$$y = \mu \mathbf{1}_n + X_\tau \tau + X_\beta \beta + e \qquad (۲.۲)$$

که $\mathbf{1}_n$ یک بردار ستونی با مرتبه‌ی n که همه‌ی درایه‌های آن، برابر واحد می‌باشد و X_β ماتریس وقوع بلوک به صورت زیر است:

$$X_\beta = diag(\mathbf{1}_{k_1}, \mathbf{1}_{k_2}, \ldots, \mathbf{1}_{k_b})$$

و $X_\tau = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_v)$ ماتریس وقوع تیماری می‌باشد که $\mathbf{x}_i$ یک بردار ستونی با r_i درایه‌ی یک و $(n - r_i)$ درایه‌ی صفر می‌باشد به قسمی که $\mathbf{x}_i'\mathbf{x}_i = r_i$ و $\mathbf{x}_i'\mathbf{x}_{i'} = 0$ برای مقادیر $i \neq i' = 1, 2, \ldots, v$ می باشد. همان‌گونه که ملاحظه می‌شود بردار $\mathbf{x}_i$، دارای n مؤلفه می‌باشد.

۲–۲–۲ معادلات نرمال و نرمال کاهش یافته

برای برآورد پارامترهای مدل روش حداقل مربعات معمولی را به کار می‌گیریم، یعنی مجموع مربعات خطا را نسبت به پارامترها مینیمم می‌کنیم. معادلاتی که با روش مشتق گیری جزئی نسبت به هریک از پارامترها

حاصل می شوند را معادلات نرمال می‌نامند.

معادلات نرمال برای μ و τ_i و β_j به صورت زیر می‌باشند:

$$n\widehat{\mu} + \sum_{i=1}^{v} r_i\widehat{\tau_i} + \sum_{j=1}^{b} k_j\widehat{\beta_j} = G \qquad (۳.۲)$$

$$r_i\widehat{\mu} + r_i\widehat{\tau_i} + \sum_{j=1}^{b} n_{ij}\widehat{\beta_j} = T_i \quad i = 1, 2, \ldots, v \qquad (۴.۲)$$

$$k_j\widehat{\mu} + \sum_{i=1}^{v} n_{ij}\widehat{\tau_i} + k_j\widehat{\beta_j} = B_j \quad j = 1, 2, \ldots, b \qquad (۵.۲)$$

که $T_i = \sum_{i=1}^{v}\sum_{j=1}^{b} y_{ijl}$ مجموع مشاهدات مربوط به تیمار i ام و $B_j = \sum_{i=1}^{v}\sum_{l=1}^{n_{ij}} y_{ijl}$ مجموع مشاهدات مربوط به بلوک j ام می باشد و $G = \sum_i T_i = \sum_j B_j$ مجموع کل مشاهدات می‌باشد. همچنین $\widehat{\mu}$، $\widehat{\tau_i}$ و $\widehat{\beta_j}$ به ترتیب برآورد پارامتر مشترک برای تمام تیمارها(میانگین کل)، برآورد اثر i ـامین تیمار و برآورد اثر j ـامین بلوک می‌باشند.

رابطه‌های (۳.۲)، (۴.۲) و (۵.۲) را می‌توان به فرم ماتریسی زیر نوشت:

$$\begin{pmatrix} \mathbf{1}_n'\mathbf{1}_n & \mathbf{1}_n'X_\tau & \mathbf{1}_n'X_\beta \\ X_\tau'\mathbf{1}_n & X_\tau'X_\tau & X_\tau'X_\beta \\ X_\beta'\mathbf{1}_n & X_\beta'X_\tau & X_\beta'X_\beta \end{pmatrix} \begin{pmatrix} \widehat{\mu} \\ \widehat{\tau} \\ \widehat{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{1}_n'y \\ X_\tau'y \\ X_\beta'y \end{pmatrix} \qquad (۶.۲)$$

که با توجه به خواص $\mathbf{1}$ و X_τ و X_β رابطه‌ی (۶.۲) را می‌توان به‌صورت زیر بازنویسی کرد:

$$\begin{pmatrix} \mathbf{1}_n'\mathbf{1}_n & \mathbf{1}_v'R & \mathbf{1}_b'K \\ R\mathbf{1}_v & R & N \\ K\mathbf{1}_b & N' & K \end{pmatrix} \begin{pmatrix} \widehat{\mu} \\ \widehat{\tau} \\ \widehat{\beta} \end{pmatrix} = \begin{pmatrix} G \\ T \\ B \end{pmatrix} \qquad (۷.۲)$$

به‌طوری که ماتریس‌های R و K ماتریس‌های قطری به ترتیب از r_i ها و k_j ها می‌باشند. همچنین برای هر طرح بلوکی، ماتریس $N = (n_{ij})$ با مرتبه‌ی $v \times b$ وجود دارد که آن را ماتریس وقوع می‌گویند. ضمناً بردارهای زیر را در نظر می‌گیریم:

$$\begin{aligned} T' &= (T_1, T_2, \ldots, T_v) \\ B' &= (B_1, B_2, \ldots, B_b) \\ \tau' &= (\tau_1, \tau_2, \ldots, \tau_v) \\ \beta' &= (\beta_1, \beta_2, \ldots, \beta_b) \end{aligned}$$

با استفاده از سومین معادله‌ی رابطه‌ی (۷.۲) داریم: