۵.۱.۵ خروجی کامپیوتری نمونه‌ای

در شکل ۶.۵ خروجی کامپیوتری برای آزمون داده‌های میزان سختی مثال ۱.۵ را بر اساس شیوهٔ مدل کلی خطی (GLM) در نرم‌افزار SAS نشان داده‌ایم. خاطرنشان می‌کنیم که تحلیل اصلی در جدول ۵.۵ با استفاده از داده‌های کدگذاری شده انجام گرفته است. (با کم کردن عدد ۹٫۵ از پاسخهای خام و ضرب نتیجه در عدد ۱۰ مقادیر خام کدگذاری شده‌اند). تحلیل کامپیوتری براساس پاسخهای خام بوده است. در نتیجه، مجموع مربعات در شکل ۶.۵ برابر مقادیر مندرج در جدول ۵.۵ هستند که بر عدد ۱۰۰ تقسیم شده‌اند.

مجموع مربعات مدل که در تحلیل واریانس گزارش شده شامل مجموع مربعات تیماری به‌علاوهٔ مجموع مربعات بلوکی است؛ یعنی،

$$\begin{aligned} SS_{\text{مدل}} &= SS_{\text{نوع تیغه}} + SS_{\text{بلوک}} \\ &= ۰٫۳۸۵۰ + ۰٫۸۲۵۰ = ۱٫۲۱ \end{aligned}$$

در خروجی کامپیوتر منابع تغییرات جداگانه مربوط به‌انواع تیغه‌ها و بلوکها نیز مشخص شده‌اند. توجه کنید که مجموع مربعات نوع I و III برای این دو عامل یکسان‌اند، که همیشه برای داده‌های متعادل چنین است.

R^2 برای این مدل که به‌صورت،

$$R^2 = \frac{SS_{\text{مدل}}}{SS_{\text{خطا}}} = \frac{۱٫۲۱}{۱٫۲۹} = ۰٫۹۳۷۹۸۴$$

محاسبه می‌شود نشان می‌دهد که نزدیک به ۹۴ درصد تغییرپذیری داده‌ها به‌وسیلهٔ مدل (یعنی، عوامل نوع تیغه‌ها و بلوکها) توجیه می‌شود. این موضوع همراه با تحلیل رضایتبخش مانده‌ای، حاکی از آن‌است که مدل خیلی خوب به‌داده‌ها برازانده شده است.

۲.۵ طرح مربع لاتین

در بخش ۱.۵ طرح بلوکی کامل تصادفی‌شده را به‌عنوان طرحی برای تقلیل خطای مانده در یک آزمایش با حذف تغییرپذیری ناشی از متغیر اغتشاش معلوم و قابل کنترل معرفی کردیم. انواع دیگر طرحهایی وجود دارند که از اصل بلوکبندی استفاده می‌کنند. مثلاً، گیریم آزمایشگری اثرهای پنج فرمولبندی مختلف یک ترکیب قابل انفجار را که در ساخت دینامیت استفاده می‌شود با مشاهدهٔ نیروی انفجار مطالعه می‌کند. هر فرمولبندی از یک دسته مواد خام که تنها برای پنج فرمولبندی مورد آزمون تکافو می‌کند تهیه می‌شود. به‌علاوه فرمولبندیها به‌وسیلهٔ افراد مختلفی آماده می‌شوند که به‌دلیل اختلاف مهارت و تجربهٔ آنها ممکن است تفاوت اساسی داشته باشند. پس، به‌نظر می‌رسد

جدول ۹.۵ طرح مربع لاتین برای مسألهٔ فرمولبندی دینامیت

دسته‌های مواد خام	عملگرها ۱	۲	۳	۴	۵
۱	$A=۲۴$	$B=۲۰$	$C=۱۹$	$D=۲۴$	$E=۲۴$
۲	$B=۱۷$	$C=۲۴$	$D=۳۰$	$E=۲۷$	$A=۳۶$
۳	$C=۱۸$	$D=۳۸$	$E=۲۶$	$A=۲۷$	$B=۲۱$
۴	$D=۲۶$	$E=۳۱$	$A=۲۶$	$B=۲۳$	$C=۲۲$
۵	$E=۲۲$	$A=۳۰$	$B=۲۰$	$C=۲۹$	$D=۳۱$

دو عامل اغتشاش وجود دارند که باید «به‌طور متوسط از طرح خارج شوند»؛ دسته‌های مواد خام و عملگرها. طرح مناسب برای این مسأله عبارت از آزمون هر فرمولبندی دقیقاً یک‌بار از هر دستهٔ مواد خام و تهیهٔ هر فرمولبندی دقیقاً یک‌بار به‌وسیلهٔ هر یک از پنج عملگر است. طرح حاصل را که در جدول ۹.۵ نشان داده‌ایم، طرح مربع لاتین می‌نامند. توجه کنید که طرح، یک آرایش مربع است و پنج فرمولبندی (یا تیمارها) را با حروف لاتین A, B, C, D, E نشان داده‌ایم، لذا آن‌را مربع لاتین نامیده‌ایم. می‌بینیم که هر دو دسته‌های مواد خام (سطرها) و عملگرها (ستونها) بر تیمارها متعامدند.

طرح مربع لاتین را برای حذف دو منبع اغتشاش تغییرپذیری به‌کار می‌برند، یعنی طرحی که به‌صورت سیستماتیک، بلوکبندی در دو جهت را میسر می‌سازد. پس، در واقع سطرها و ستونها معرف دو قید برای تصادفی شدن‌اند. به‌طور کلی، مربع لاتین برای p عامل یا یک مربع لاتین $p \times p$، مربعی شامل p سطر و p ستون است. هر یک از $p^۲$ خانهٔ حاصل، شامل یکی از p حرف متناظر با تیمارهاست، و هر حرف یک‌بار و تنها یک‌بار در هر سطر و ستون رخ می‌دهد. چند مثال برای مربع لاتین عبارت‌اند از:

۴ × ۴	۵ × ۵	۶ × ۶
$ABDC$	$ADBEC$	$ADCEBF$
$BCAD$	$DACBE$	$BAECFD$
$CDBA$	$CBEDA$	$CEDFAB$
$DACB$	$BEACD$	$DCFBEA$
	$ECDAB$	$FBADCE$
		$EFBADC$

مدل آماری برای مربع لاتین عبارت است از:

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk} \begin{cases} i = ۱, ۲, \cdots, p \\ j = ۱, ۲, \cdots, p \\ k = ۱, ۲, \cdots, p \end{cases} \qquad (۲۲.۵)$$

که در آن y_{ijk} مشاهدهٔ jامین تیمار در سطر iام و ستون kام، μ میانگین کل، α_i اثر سطر iام، τ_j اثر تیمار jام، β_k اثر ستون kام، و ϵ_{ijk} خطای تصادفی است. مدل کاملاً جمعی است، یعنی هیچ اثر متقابلی بین سطرها، ستونها، و تیمارها وجود ندارد. چون در هر خانه تنها یک مشاهده وجود دارد، لذا برای نشان دادن یک مشاهدهٔ خاص، تنها دو زیرنویس از سه زیرنویس i، j، k کافی‌اند. مثلاً با مراجعه به‌مسألهٔ فرمولبندی دینامیت در جدول ۹.۵، اگر $i = ۲$ و $k = ۳$ باشد خودبه‌خود (فرمولبندی D) $j = ۴$ را پیدا می‌کنیم، و اگر $i = ۱$ و $j = ۳$ (فرمولبندی C) باشد، $k = ۳$ را به‌دست می‌آوریم. این، پیامد ظاهر شدن هر تیمار دقیقاً یک‌بار در هر سطر و ستون است.

تحلیل واریانس مرکب از افراز کل مجموع مربعات $N = p^۲$ مشاهده به‌مؤلفه‌های سطرها، ستونها، تیمارها، و خطاست، یعنی

$$SS_{\text{کل}} = SS_{\text{سطرها}} + SS_{\text{ستونها}} + SS_{\text{تیمارها}} + SS_{\text{خطا}} \qquad (۲۳.۵)$$

به‌ترتیب با درجات آزادی

$$p^۲ - ۱ = p - ۱ + p - ۱ + p - ۱ + (p - ۲)(p - ۱)$$

تحت پذیرهٔ معمول $\epsilon_{ijk} \sim \mathrm{NID}(۰, \sigma^۲)$، هر یک از مجموع مربعات طرف راست برابری (۲۳.۵) وقتی بر $\sigma^۲$ تقسیم می‌شود متغیر تصادفی مستقل با توزیع خی دو است. آمارهٔ مناسب برای آزمون متفاوت نبودن میانگینهای تیماری،

$$F_۰ = \frac{MS_{\text{تیمارها}}}{MS_{\text{خطا}}}$$

است که تحت فرض صفر، توزیع $F_{p-۱,(p-۲)(p-۱)}$ دارد. می‌توان آزمون نبودن اثر سطری و نبودن اثر ستونی را با تشکیل نسبت $MS_{\text{سطرها}}$ یا $MS_{\text{ستونها}}$ به $MS_{\text{خطا}}$ انجام داد. اما، چون سطرها و ستونها معرف محدودیتهایی در تصادفی بودن‌اند لذا این آزمونها نمی‌توانند مناسب باشند.

شیوهٔ محاسباتی تحلیل واریانس را در جدول ۱۰.۵ نشان داده‌ایم. بنابر فرمولهای محاسباتی برای مجموع مربعات، می‌بینیم که تحلیل، توسیع ساده‌ای از طرح بلوکی تصادفی‌شده است، با مجموع مربعات سطرها که از جمع سطرها به‌دست می‌آید.

جدول ۱۰.۵ تحلیل واریانس برای طرح مربع لاتین

منبع تغییر	مجموع مربعات	درجات آزادی	میانگین مربعات	$F_\circ$
تیمارها	$SS_{\text{تیمارها}} = \sum_{j=1}^{p} \frac{y_{.j.}^2}{p} - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{تیمارها}}}{p-1}$	$F_\circ = \frac{MS_{\text{تیمارها}}}{MS_{\text{خطا}}}$
سطرها	$SS_{\text{سطرها}} = \sum_{i=1}^{p} \frac{y_{i..}^2}{p} - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{سطرها}}}{p-1}$	
ستونها	$SS_{\text{ستونها}} = \sum_{k=1}^{p} \frac{y_{..k}^2}{p} - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{ستونها}}}{p-1}$	
خطا	$SS_{\text{خطا}}$(با تفریق)	$(p-2)(p-1)$	$\frac{SS_{\text{خطا}}}{(p-2)(p-1)}$	
کل	$SS_{\text{کل}} = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$	p^2-1		

مثال ۴.۵

مسألهٔ فرمولبندی دینامیت را که قبلاً شرح آن گذشت در نظر بگیرید، که در آن هر دو دسته‌های مواد خام و عملگرها محدودیت تصادفی بودن را نشان می‌دهند. طرح این آزمایش را که در جدول ۹.۵ نشان داده‌ایم یک مربع لاتین ۵ × ۵ است. بعد از کدگذاری داده‌ها با کم کردن ۲۵ از هر مشاهده، داده‌های جدول ۱۱.۵ را به‌دست آورده‌ایم. مجموع مربعات کل، دسته‌ها (سطرها)، و

جدول ۱۱.۵ داده‌های کدگذاری شده برای مسألهٔ فرمولبندی دینامیت

دسته‌های مواد خام	عملگرها ۱	۲	۳	۴	۵	$y_{i..}$
۱	$A=-1$	$B=-5$	$C=-6$	$D=-1$	$E=-1$	-14
۲	$B=-8$	$C=-1$	$D=5$	$E=2$	$A=11$	۹
۳	$C=-7$	$D=13$	$E=1$	$A=2$	$B=-4$	۵
۴	$D=1$	$E=6$	$A=1$	$B=-2$	$C=-3$	۳
۵	$E=-3$	$A=5$	$B=-5$	$C=4$	$D=6$	۷
$y_{..k}$	-18	۱۸	-4	۵	۹	$10 = y_{...}$

عملگرها (ستونها) را به‌صورت زیر محاسبه کرده‌ایم:

$$SS_{\text{کل}} = \sum_i \sum_j \sum_k y_{ijk}^{۲} - \frac{y_{...}^{۲}}{N}$$

$$= ۶۸۰ - \frac{(۱۰)^{۲}}{۲۵} = ۶۷۶٫۰۰$$

$$SS_{\text{دسته‌ها}} = \sum_{i=۱}^{p} \frac{y_{i..}^{۲}}{p} - \frac{y_{...}^{۲}}{N}$$

$$= \frac{(-۱۴)^{۲} + ۹^{۲} + ۵^{۲} + ۳^{۲} + ۷^{۲}}{۵} - \frac{(۱۰)^{۲}}{۲۵} = ۶۸٫۰۰$$

$$SS_{\text{عملگرها}} = \sum_{k=۱}^{p} \frac{y_{..k}^{۲}}{p} - \frac{y_{...}^{۲}}{N}$$

$$= \frac{(-۱۸)^{۲} + ۱۸^{۲} + (-۴)^{۲} + ۵^{۲} + ۹^{۲}}{۵} - \frac{(۱۰)^{۲}}{۲۵} = ۱۵۰٫۰۰$$

مجموعهای مربوط به‌تیمارها (حروف لاتین) عبارت‌اند از

حرف لاتین	مجموع تیمار
A	$y_{.۱.} = ۱۸$
B	$y_{.۲.} = -۲۴$
C	$y_{.۳.} = -۱۳$
D	$y_{.۴.} = ۲۴$
E	$y_{.۵.} = ۵$

مجموع مربعات حاصل از فرمولبندیها از روی این مجموعها محاسبه می‌شود

$$SS_{\text{فرمولبندیها}} = \sum_{j=۱}^{p} \frac{y_{.j.}^{۲}}{p} - \frac{y_{...}^{۲}}{N}$$

$$= \frac{۱۸^{۲} + (-۲۴)^{۲} + (-۱۳)^{۲} + ۲۴^{۲} + ۵^{۲}}{۵} - \frac{(۱۰)^{۲}}{۲۵} = ۳۳۰٫۰۰$$

مجموع مربعات خطا با تفریق به‌دست می‌آید:

$$SS_{\text{خطا}} = SS_{\text{کل}} - SS_{\text{دسته‌ها}} - SS_{\text{عملگرها}} - SS_{\text{فرمولبندیها}}$$

$$= ۶۷۶٫۰۰ - ۶۸٫۰۰ - ۱۵۰٫۰۰ - ۳۳۰٫۰۰ = ۱۲۸٫۰۰$$

تحلیل واریانس در جدول ۱۲.۵ خلاصه شده است. نتیجه می‌شود که تفاوتی معنی‌دار در

جدول ۱۲.۵ تحلیل واریانس برای آزمایش فرمولبندی دینامیت

منبع تغییر	مجموع مربعات	درجات آزادی	میانگین مربعات	F_0
فرمولبندیها	۳۳۰٫۰۰	۴	۸۲٫۵۰	۷٫۷۳*
دسته‌های مواد خام	۶۸٫۰۰	۴	۱۷٫۰۰	
عملگرها	۱۵۰٫۰۰	۴	۳۷٫۵۰	
خطا	۱۲۸٫۰۰	۱۲	۱۰٫۶۷	
کل	۶۷۶٫۰۰	۲۴		

* در یک درصد معنی‌دار.

میانگینهای نیروی انفجار حاصل از فرمولبندیهای مختلف دینامیت وجود داشته است. همچنین، معلوم می‌شود که بین عملگرها تفاوتهایی موجود است. لذا بلوکبندی این عامل، پیش احتیاط خوبی بوده است. گواهی قوی از وجود اختلاف بین دسته‌های مواد خام موجود نیست، و بنابراین به‌نظر می‌رسد که در این آزمایش خاص، بدون آنکه لازم باشد نگران این منبع تغییرپذیری بوده‌ایم. اما، معمولاً بلوکبندی دسته‌های مواد خام فکر خوبی است. ■

مانند هر مسألهٔ طرح، آزمایشگر باید کفایت مدل را به‌وسیلهٔ بررسی و رسم مانده‌ها رسیدگی کند. برای یک مربع لاتین، مانده‌ها به‌وسیلهٔ

$$\begin{aligned} e_{ijk} &= y_{ijk} - \hat{y}_{ijk} \\ &= y_{ijk} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..k} + ۲\bar{y}_{...} \end{aligned}$$

به‌دست می‌آیند. خواننده باید مانده‌ها را برای مثال ۴.۵ به‌دست آورده و نمودارهای مناسب را رسم کند.

مربع لاتینی را که سطر اول و ستون اول آن شامل حروفی با ترتیب الفباست مربع لاتین استاندارد می‌نامیم. طرح مثال ۴.۵ یک مربع لاتین استاندارد است. مربع لاتین استاندارد را همیشه می‌توان با نوشتن اولین سطر به‌ترتیب الفبا، و سپس نوشتن هر یک از سطرهای بعدی به‌شکل سطر بالای آن‌که یک مکان به‌طرف چپ منتقل شده باشد به‌دست آورد. در جدول ۱۳.۵ چند واقعیت مهم دربارهٔ مربعات لاتین و مربعات لاتین استاندارد خلاصه شده است.

مانند هر طرح آزمایش دیگر، در طرح مربع لاتین مشاهدات باید به‌ترتیب تصادفی انتخاب شوند. شیوهٔ درست تصادفی‌کردن آن است که مربع خاصی که به‌کار می‌رود به‌طور تصادفی انتخاب شود.

جدول ۱۳.۵ مربعات لاتین استاندارد و تعداد مربعات لاتین با اندازه‌های مختلف*

اندازه	۳ × ۳	۴ × ۴	۵ × ۵	۶ × ۶	۷ × ۷	$p \times p$
مثالهایی از مربعات استاندارد	ABC BCA CAB	$ABCD$ $BCDA$ $CDAB$ $DABC$	$ABCDE$ $BAECD$ $CDAEB$ $DEBAC$ $ECDBA$	$ABCDEF$ $BCFADE$ $CFBEAD$ $DEABFC$ $EACFCB$ $FDECBA$	$ABCDEFG$ $BCDEFGA$ $CDEFGAB$ $DEFGABC$ $EFGABCD$ $FGABCDE$ $GABCDEF$	$ABC \cdots P$ $BCD \cdots A$ $CDE \cdots B$ $\vdots$ $PAB \cdots (P-۱)$
تعداد مربعات استاندارد	۱	۴	۵۶	۹۴۰۸	۱۶۹۴۲۰۸۰	—
تعداد کل مربعات لاتین	۱۲	۵۷۶	۱۶۱۲۸۰	۸۱۸۸۵۱۲۰۰	۶۱۴۷۹۴۱۹۹۰۴۰۰۰	$p!(p-۱)!\times$ (تعداد مربعات استاندارد)

* برخی از اطلاعات این جدول از چاپ چهارم جدولهای آماری برای تحقیقات زیست‌شناسی، کشاورزی، و پزشکی که فیشر و یتس (۱۹۵۳) تدوین کرده‌اند استخراج شده است. اطلاعات کمی در مورد خواص مربعات لاتین بزرگتر از ۷ × ۷ در اختیار است.