جدول ۲.۵ تحلیل واریانس برای طرح بلوکی کامل تصادفی‌شده

منبع متغیر	مجموع مربعات	درجات آزادی	میانگین مربعات	$F_\circ$
تیمارها	$\sum_{i=1}^{a} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{N}$	$a-1$	$\frac{SS_{\text{تیمارها}}}{a-1}$	$\frac{MS_{\text{تیمارها}}}{MS_{\text{خطا}}}$
بلوکها	$\sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{N}$	$b-1$	$\frac{SS_{\text{بلوکها}}}{b-1}$	
خطا	$SS_{\text{خطا}}$ (با تفریق)	$(a-1)(b-1)$	$\frac{SS_{\text{خطا}}}{(a-1)(b-1)}$	
کل	$\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N}$	$N-1$		

به‌این دلیل، آزمون F را از جدول تحلیل واریانس کنار گذاشته‌ایم. اما، به‌عنوان شیوه‌ای تقریبی در بررسی اثر متغیر بلوکبندی، امتحان نسبت $MS_{\text{بلوکها}}$ به $MS_{\text{خطا}}$ مسلماً منطقی است. اگر این نسبت بزرگ باشد، نتیجه می‌شود که عامل بلوکبندی اثری بزرگ دارد و تقلیل اغتشاش که با بلوکبندی کردن حاصل شده است احتمالاً در اصلاح دقت مقایسهٔ میانگینهای تیماری مفید بوده است.

معمولاً شیوه را به‌اختصار در یک جدول تحلیل واریانس نظیر آنچه که در جدول ۲.۵ نشان داده‌ایم می‌آوریم. فرمولهای محاسباتی برای مجموع مربعات را می‌توان برای عناصر برابری (۷.۵) برحسب مجموعهای تیمارها و بلوکها به‌دست آورد. این فرمولهای محاسباتی عبارت‌اند از

$$SS_{\text{کل}} = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N} \tag{۹.۵}$$

$$SS_{\text{تیمارها}} = \sum_{i=1}^{a} \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{N} \tag{۱۰.۵}$$

$$SS_{\text{بلوکها}} = \sum_{j=1}^{b} \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{N} \tag{۱۱.۵}$$

و مجموع مربعات خطا با تفریق کردن به‌صورت زیر به‌دست می‌آید

$$SS_{\text{خطا}} = SS_{\text{کل}} - SS_{\text{تیمارها}} - SS_{\text{بلوکها}} \tag{۱۲.۵}$$

مثال ۱.۵

آزمایش آزمون سختی را که در بخش ۱.۵ شرح آن گذشت در نظر بگیرید. چهار تیغه و چهار قطعهٔ فلزی داریم. در نتیجهٔ طرح بلوکی کامل تصادفی‌شده، هر تیغه یک‌بار روی هر قطعه آزمون می‌شود. داده‌های حاصل را برای راحتی در جدول ۳.۵ تکرار کرده‌ایم. خاطرنشان می‌کنیم ترتیبی

که هر تیغه بر قطعه‌ای خاص آزمون شده، تصادفی بوده است. برای ساده کردن محاسبات، داده‌های اصلی را کدگذاری کرده‌ایم به‌این طریق که هر مشاهده را از عدد ۹٫۵ کم کرده و سپس آن‌را در ۱۰ ضرب نموده‌ایم. به‌این ترتیب داده‌های جدول ۴.۵ حاصل شده‌اند. مجموع مربعات به‌صورت زیر به‌دست می‌آیند.

$$SS_{\text{کل}} = \sum_{i=۱}^{۴}\sum_{j=۱}^{۴} y_{ij}^{۲} - \frac{y_{..}^{۲}}{N}$$

$$= ۱۵۴٫۰۰ - \frac{(۲۰)^{۲}}{۱۶} = ۱۲۹٫۰۰$$

$$SS_{\text{تیمارها}} = \sum_{i=۱}^{۴} \frac{y_{i.}^{۲}}{b} - \frac{y_{..}^{۲}}{N}$$

$$= \frac{(۳)^{۲} + (۴)^{۲} + (-۲)^{۲} + (۱۵)^{۲}}{۴} - \frac{(۲۰)^{۲}}{۱۶} = ۳۸٫۵۰$$

$$SS_{\text{بلوکها}} = \sum_{j=۱}^{۴} \frac{y_{.j}^{۲}}{a} - \frac{y_{..}^{۲}}{N}$$

$$= \frac{(-۴)^{۲} + (-۳)^{۲} + (۹)^{۲} + (۱۸)^{۲}}{۴} - \frac{(۲۰)^{۲}}{۱۶} = ۸۲٫۵۰$$

$$SS_{\text{خطا}} = SS_{\text{کل}} - SS_{\text{تیمارها}} - SS_{\text{بلوکها}}$$

$$= ۱۲۹٫۰۰ - ۳۸٫۵۰ - ۸۲٫۵۰ = ۸٫۰۰$$

تحلیل واریانس را در جدول ۵.۵ نشان داده‌ایم. با در نظر گرفتن $\alpha = ۰٫۰۵$، مقدار بحرانی F به‌صورت $F_{۰٫۰۵,۳,۹} = ۳٫۸۶$. چون $۳٫۸۶ < ۱۴٫۴۴$، نتیجه می‌گیریم که نوع تیغه در میانگین

جدول ۳.۵ طرح بلوکی کامل تصادفی‌شده در آزمایش آزمون میزان سختی

نوع تیغه	قطعه (بلوک) ۱	۲	۳	۴
۱	۹٫۳	۹٫۴	۹٫۶	۱۰٫۰
۲	۹٫۴	۹٫۳	۹٫۸	۹٫۹
۳	۹٫۲	۹٫۴	۹٫۵	۹٫۷
۴	۹٫۷	۹٫۶	۱۰٫۰	۱۰٫۲

جدول ۴.۵ داده‌های کدگذاری شده برای آزمایش آزمون میزان سختی

نوع تیغه	قطعه (بلوک) ۱	۲	۳	۴	$y_{i.}$
۱	۲−	۱−	۱	۵	۳
۲	۱−	۲−	۳	۴	۴
۳	۳−	۱−	۰	۲	۲−
۴	۲	۱	۵	۷	۱۵
$y_{.j}$	۴−	۳−	۹	۱۸	$۲۰ = y_{..}$

جدول ۵.۵ تحلیل واریانس برای آزمایش آزمون میزان سختی

منبع تغییر	مجموع مربعات	درجات آزادی	میانگین مربعات	F_0
تیمارها (نوع تیغه)	۳۸٫۵۰	۳	۱۲٫۸۳	۱۴٫۴۴
بلوکها (قطعات)	۸۲٫۵۰	۳	۲۷٫۵۰	
خطا	۸٫۰۰	۹	۰٫۸۹	
کل	۱۲۹٫۰۰	۱۵		

میزان سختی قرائت شده مؤثر بوده است. همچنین، قطعات (بلوکها) به‌صورتی معنی‌دار متفاوت به‌نظر می‌رسند، زیرا که میانگین مربعات برای بلوکها نسبت به خطا بزرگ است.

جالب است نتایج حاصل از همین داده‌ها را وقتی از طرحهای بلوکی تصادفی‌شده اطلاعی نداریم ببینیم. گیریم برای هر تیغه از چهار قطعه استفاده کرده باشیم، که به‌هر یک از آنها تیغه‌ای را به‌تصادف امتحان کرده‌ایم. و برحسب شانس همان نتایج جدول ۳.۵ حاصل شده باشند. تحلیل غلط این داده‌ها را به‌صورت طرح کاملاً تصادفی‌شدهٔ تک عاملی در جدول ۶.۵ نشان داده‌ایم. چون $F_{0.05,3,12} = 3.49$، فرض مساوی بودن میانگین میزان سختی چهار تیغه را نمی‌توانیم رد کنیم. پس، طرح بلوکی تصادفی‌شده برای آشکارکردن تفاوت بین چهار تیغه مقدار اغتشاش داده‌ها را به‌اندازهٔ کافی کاهش داده است. ■

جدول ۶.۵ تحلیل غلط آزمایش آزمون میزان سختی به‌صورت طرح کاملاً تصادفی‌شده

منبع تغییر	مجموع مربعات	درجات آزادی	میانگین مربعات	F_0
نوع تیغه	۳۸٫۵۰	۳	۱۲٫۸۳	۱٫۷۰
خطا	۹۰٫۵۰	۱۲	۷٫۵۴	
کل	۱۲۹٫۰۰	۱۵		

جمعی بودن مدل بلوکی تصادفی‌شده. مدلی خطی که برای طرح بلوکی تصادفی‌شده به‌کار بردیم، یعنی؛

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

کاملاً جمعی است. جمعی، به‌معنای آن‌است که مثلاً اگر اولین تیمار موجب افزایش امید ریاضی پاسخ به‌اندازهٔ ۵ واحد شود ($\tau_1 = ۵$) و اگر اولین بلوک امید ریاضی پاسخ را به‌اندازهٔ ۲ واحد افزایش دهد ($\beta_1 = ۲$)، آن‌گاه برای تیمار ۱ و بلوک ۱ با هم امید ریاضی پاسخ به‌اندازهٔ $E(y_{11}) = \mu + \tau_1 + \beta_1 = \mu + ۵ + ۲ = \mu + ۷$ واحد افزایش پیدا می‌کند. به‌طورکلی، تیمار ۱ همیشه امید ریاضی پاسخ را به‌اندازهٔ ۵ واحد بیشتر از مجموع میانگین کل و اثر بلوکی نشان می‌دهد.

هرچند چنین مدل سادهٔ جمعی غالباً مفید است، اما وضعیتهایی وجود دارند که برای آنها این مدل مناسب نیست. مثلاً، گیریم چهار فرمولبندی یک فراوردهٔ شیمیایی را با استفاده از شش دسته مواد خام مقایسه می‌کنیم؛ دسته‌های مواد خام را به‌عنوان بلوکها در نظر می‌گیریم. اگر ناخالصی در دستهٔ ۲ اثر معکوس در فرمولبندی ۲ داشته باشد و به‌صورتی غیرمعمول موجب افت محصول شود، اما در فرمولبندیهای دیگر مؤثر نباشد، آن‌گاه اثر متقابل بین فرمولبندیها (یا تیمارها) و دسته‌ها (یا بلوکها) وجود دارد. همچنین اثرهای متقابل بین تیمارها و بلوکها وقتی پاسخ با مقیاس غلط اندازه‌گیری شده باشد، می‌توانند رخ دهند. پس، رابطه‌ای ضربی بین واحدهای اصلی مثل،

$$E(y_{ij}) = \mu\tau_i\beta_j$$

برقرار است، که به‌مقیاس لگاریتمی، خطی یا جمعی است، زیرا مثلاً

$$\ln E(y_{ij}) = \ln \mu + \ln \tau_i + \ln \beta_j$$

یا

$$E(y^*_{ij}) = \mu^* + \tau^*_i + \beta^*_j$$

است و از جدول VII پیوست با $(a-۱)(b-۱)$ درجهٔ آزادی برای تعیین دامنه‌های معنی‌دار بودن استفاده می‌کنیم.

مثال ۲.۵

چون در مثال ۱.۵ نتیجه گرفتیم که میانگین سختی از تیغه‌ای به تیغهٔ دیگر تفاوت می‌کند، لذا علاقه‌مندیم که ماهیت خاص این تفاوتها را کشف کنیم. میانگینهای تیماری به‌ترتیب صعودی به‌صورت زیرند:

$$\bar{y}_{۳.} = -۰٫۵۰ \qquad \bar{y}_{۱.} = ۰٫۷۵ \qquad \bar{y}_{۲.} = ۱٫۰۰ \qquad \bar{y}_{۴.} = ۳٫۷۵$$

و خطای معیار میانگین تیماری برابر است با؛

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_{\text{خطا}}}{b}} = \sqrt{\frac{۰٫۸۹}{۴}} = ۰٫۴۷$$

از جدول VII پیوست دامنه‌های معنی‌دار زیر را به‌دست می‌آوریم

$$r_{۰٫۰۵}(۲,۹) = ۳٫۲۰ \qquad r_{۰٫۰۵}(۳,۹) = ۳٫۳۴ \quad \text{و} \quad r_{۰٫۰۵}(۴,۹) = ۳٫۴۱$$

پس، کمترین دامنه‌های معنی‌دار عبارت‌اند از

$$R_۲ = r_{۰٫۰۵}(۲,۹)S_{\bar{y}_{i.}} = (۳٫۲۰)(۰٫۴۷) = ۱٫۵۰$$
$$R_۳ = r_{۰٫۰۵}(۳,۹)S_{\bar{y}_{i.}} = (۳٫۳۴)(۰٫۴۷) = ۱٫۵۷$$
$$R_۴ = r_{۰٫۰۵}(۴,۹)S_{\bar{y}_{i.}} = (۳٫۴۱)(۰٫۴۷) = ۱٫۶۰$$

و مقایسه‌های دوبه‌دو عبارت‌اند از

$$۴ \text{ با } ۳ : ۳٫۷۵ - (-۰٫۵۰) = ۴٫۲۵ > ۱٫۶۰(R_۴)$$
$$۴ \text{ با } ۱ : ۳٫۷۵ - ۰٫۷۵ = ۳٫۰۰ > ۱٫۵۷(R_۳)$$
$$۴ \text{ با } ۲ : ۳٫۷۵ - ۱٫۰۰ = ۲٫۷۵ > ۱٫۵۰(R_۲)$$
$$۲ \text{ با } ۳ : ۱٫۰۰ - (-۰٫۵۰) = ۱٫۵۰ < ۱٫۵۷(R_۳)$$

توجه کنید که چون مقایسهٔ $\bar{y}_{۲.}$ و $\bar{y}_{۳.}$ بقیهٔ میانگینها را در برمی‌گیرد و مقایسه نشان می‌دهد که فرض $H_۰ : \mu_۲ = \mu_۳$ را نمی‌توان رد کرد، لذا بلافاصله نتیجه می‌گیریم که دیگر جفتهای میانگینها