

۱۳۰۳/۷/۱۷

روش بارمتری ⇐ توزیع جامعه مشخص است.
روش نامبارمتری ⇐ توزیع جامعه مشخص نیست.

حزب آزمایش های تک عامل: باید که چون دسته ها جامعه را با هم ضایع می کنیم میخواهیم تک تک اینها را

خطا اشتراک است.
تفاوت مدل در طرح آزمایش ها و رگرسیون.
رگرسیون } متغیرهای کیفی ⇐ X کیفی
 } متغیرهای پانچ ⇐ Y کیفی
طرح } عامل ⇐ X کیفی
 } پاسخ ⇐ Y کیفی

هدف در طرح رگرسیون متفاوت است. در رگرسیون هدف رابطه بین X و Y و در طرح هدف مقایسه X ها است و پیدا کردن بهترین X. در رگرسیون هدف مدل است.

متغیر کیفی را عامل می گویند. Y در هر صورت کیفی است.

مثال: رگرسیون ⇐ X هزینه تبلیغات ⇐ کیفی
 Y میزان فروش ⇐ کیفی
مثال طرح: X نوع تبلیغات
 Y میزان فروش
۱- تلویزیون
۲- رسانه مجازی
۳- روزنامه
۴- میزبان فروش

اگر Y کیفی نبود در طرح می شود گسسته
اشتراک رگرسیون و طرح ⇐ متغیر پاسخ (Y) کیفی است.

اصول انجام آزمایش: ① تصادفی باشند (خطای سیستماتیک را کم می کنیم) ② تکرار (هر چه تکرار بیشتر باشد یعنی تعداد بیشتر باشد) باعث

کاهش خطای می شود. ③ بلوک بندی کردن (بلوک بندی باعث کاهش خطای می شود)

طرح آزمایش ها: آزمایش را طوری طرح ریزی کردیم که بتوانیم به یکی از روش های طراحی تحلیل کرد. PAPCO

روش انجاء یا طرح یک آزمایش :

- ۱- شناسایی و بیان مسأله ۲- انتخاب عوامل و سطوح ۳- انتخاب تغییرات یا سطح ۴- انتخاب طرح آزمایش ۵- انجام آزمایش
- ۶- تحلیل داده ها ۷- تصمیم گیری

« فصل سوم »

طرح های تک عامل : هدف بررسی سطوح یک عامل است .

اولین طرح : طرح تعادلی ساده : فعال : چهار نوع دارند داریم
 ۱- اگر بپذیرد تعادلی با هم دارند : این طرح برای حصول هدف کلی
 ۲- کدام بزرگتر است ؟
 بپذیرد : A¹ B² C³ D⁴

- هدفی شرایط بسیار برای بررسی کت ها

برداشت محصول : برای نوع لذر ۱ ۲ ۳ و ۴ = I
 برای تک لذر ۱ ۲ ۳ و ۴ = J
 از ۱۶ با سطح برابر اگر در نظر داشته باشیم

تعاریف مهم در طرح : عامل (factor) : آن چیزی که هدف اصلی حاصل است عامل : پذیرند

تعداد (trea) : سطوح عامل را شمار می گوئیم . تعداد : پذیرند A¹ B² C³ D⁴ تعداد : پذیرند

واحد آزمایش (unit) : آن چیزی که تیارها را آن اعمال می شوند واحد آزمایش : کت ها

یا سطح : آن چیزی که از واحدهای آزمایشی حاصل می شود . یا سطح : فزالت محصول که از هر کت حاصل می شود

تعریف طرح تعادلی ساده : ۱- یک طرح تک عاملی است (یعنی هدف فقط بررسی سطوح یک عامل است)

۷) لقبه شرایط یکسان. تنها نوع تیمارها فنلاند است. ۳) تیمارها اگر برای هر تیمار برابر باشند، طرح متعادل می گوئیم. اثر تیمارها برابر

نباشند. طرح نامتعادل می گوئیم (طرح متعادل بهتر از نامتعادل است) ۴) تیمارها باید به تصادف روی واحدهای آزمایشی اعمال

شوند. ۵) طرح تصادفی ساده ایده آل ترین طرح است اگر نتوانیم آن را اجرا کنیم.

۱۴۳۱۷۱۸

بالاتر ها را در جدول بالیج ها خلاصه می کنیم.

تیمارها

$i \backslash j$	۱	۲	...	a
۱	y_{11}	y_{12}		y_{a1}
۲	y_{12}	y_{22}		y_{a2}
...	$\vdots$	$\vdots$		$\vdots$
n	y_{1n}	y_{2n}		y_{an}
	μ_1			μ_a

a تا تیمار داریم

یک سطر داریم (فصل برداریم) که a تا سطوح دارد و برای هر کدام تا n

$i = 1, 2, \dots, a$

$j = 1, 2, \dots, n$

تکرار داریم.

میانگین تیمار a

μ_i

خطاهای آزمایشی

$$y_{ij} = \mu + T_i + \epsilon_{ij}$$

مدل طرح تصادفی ساده ۸

اثر تیمار را T_i می بینیم کل میانگین کل بالیج تیمار را μ می بینیم (تیمارها را به هم میزنیم)

خطاهای آزمایشی ۸ هر عامل ناشناخته ای که روی عامل ها تاثیر گذارد و شناسایی نگردانیم آنرا.

تیمارها باید به تصادف باشند به غیر از خطاهای سیستماتیک خطاهای آزمایشی تصادفی است.

اثر تیمار را $T_i = \mu_i - \mu$ اختلاف میانگین آن تیمار را از میانگین کل حساب می کنیم.

$$\sum_{i=1}^n T_i = 0$$

T_i برای هر تیمار ثابت است.

Subject: $S^2 = \sum \frac{(x_i - \bar{x})^2}{n-1}$
Date

شرایط مدل: خطاهای مدل باید دارای شرایط زیر باشند: 8. زیاده‌ها یا صفر باشند و در اینجواب ثابت - توانایی مدل

① $\epsilon_{ij} \sim N(0, \sigma^2)$

② $\sum_{i=1}^a T_i = 0$

یا فرض اینکه مدل با اثرات ثابت باشند (اثر تیارها ثابت باشند)

$\sum T_i = 0$ پس یا T_i ها همه باید صفر باشند یا یک سری مثبت و یک سری منفی باشند.

فرضیات (یعنی فرضیه‌ها فرض H_0 و H_1):
 $\begin{cases} H_0: \mu_1 = \mu_2 = \dots = \mu_a = \mu \\ H_1: \text{Not } H_0 \end{cases}$
 یعنی دو تا μ برداریم که برابر نباشند

یعنی تفاوت معناداری بین تیارها وجود ندارد
 $\begin{cases} H_0: T_i = 0 \quad \forall i \Rightarrow \\ H_1: \text{Not } H_0 \Rightarrow \end{cases}$
 یعنی حداقل یک وجود داشته باشد که $T_i \neq 0$ نباشد

برای آزمون کردن فرضیه فوق، آماره واریانس انتخاب می‌دهیم: Analysis of Variance (ANOVA)

آماره واریانس: یعنی تجزیه واریانس کل به مولفه‌های موجود در مدل

تیار واریانس کل

$$SST_0 = SST + SSE$$

مجموع مربعات خطاها + مجموع مربعات تیار

$$SST_0 = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$$

* جای هر اندیس را با اندیس دیگری عوض شده *

میانگین باطنی تیارها

$$\bar{y}_{i.} = \frac{y_{i.}}{n}$$

$$\bar{y}_{i.} = \sum_{j=1}^n y_{ij} \quad \leftarrow \text{حالت } \bar{y}_{i.} \text{ باشد}$$

یعنی باطنی ها را جمع کرده است

$$\bar{y}_{.j} = \frac{y_{.j}}{a}$$

$$\bar{y}_{.j} = \sum_{i=1}^a y_{ij} \quad \leftarrow \text{زیر } \bar{y}_{.j}$$

یعنی باطنی ها را جمع کرده است

$$\bar{y}_{..} = \frac{y_{..}}{an}$$

$$\bar{y}_{..} = \sum_{i=1}^a \sum_{j=1}^n y_{ij} \quad \leftarrow$$

یعنی جمع کل میانگین ها

نحوه تجزیه واریانس در طرح تصادفی ساده

$$SSTO = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.} + \bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 + \sum_{i=1}^a \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2 +$$

یک میانگین تیار کم و اضافه کردیم

$$2 \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})(\bar{y}_{i.} - \bar{y}_{..}) = \sum_{i=1}^a \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$$

اثر تیار و واریانس تیارها

فصل محاسباتی

$$SSTO = \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$N = an$$

$$SST = \sum_{i=1}^a \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N} \quad \leftarrow \text{جمع کل}$$

$$SSE = SSTO - SST = \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \sum_{i=1}^a \frac{y_{i.}^2}{n}$$

نکته: هیچ کدام از SS ها نباید منفی بدست آید.

$$SSTO = SST + SSE$$

$$df: N-1 = (a-1) + (N-a)$$

$$\downarrow$$

$$an$$

$$\downarrow$$

$$a(n-1)$$

$$MST = \frac{SST}{a-1}$$

$$F = \frac{MST}{MSE}$$

$$MSE = \frac{SSE}{N-a}$$

۱۴۰۳ / ۷ / ۱۱

شنبه

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$MS = SS$$

آماره واریانس: تجزیه واریانس کل به مولفه های موجود در مدل

امید ریاضی MS ها در مدل با اثرات تبیین شده

نتیجه می گیریم که MSE یک برآورده نادرست برای σ^2 (واریانس خطاها) است.

$$E(MSE) = \sigma^2 \Rightarrow$$

$$E(MST) = \sigma^2 + n \frac{\sum_{i=1}^n T_i^2}{a-1}$$

استفاده از امید ریاضی MS ها

$$H_0: T_i = 0 \quad \forall i$$

$$E(MST) \xrightarrow{H_0 \text{ فرض}} E(MST) = \sigma^2 = E(MSE) \quad \text{نتیجه}$$

اگر تفاوت معناداری بین تیمارها وجود نداشته باشد، تقریباً $T_i = 0$ باید باشد. یعنی $E(MST) = E(MSE)$ و MST و MSE را با هم

$$F_0 = \frac{MST}{MSE}$$

مقایسه می کنیم ← آماره آزمون

$$\frac{SST_0}{\sigma^2} \sim \chi^2_{N-1}$$

توزیع آماره آزمون: می دانیم

$$\left. \begin{array}{l} \frac{SST}{\sigma^2} \sim \chi^2_{(a-1)} \\ \frac{SSE}{\sigma^2} \sim \chi^2_{(N-a)} \end{array} \right\} \begin{array}{l} \text{مستقل از هم طبق} \\ \text{فرضه کلاسیک} \end{array} \Rightarrow F_0 = \frac{\frac{SST}{\sigma^2}}{\frac{SSE}{\sigma^2}} = \frac{(a-1)}{(N-a)} F(a-1, N-a)$$

$$F_0 = \frac{MST}{MSE} \xrightarrow{H_0} F(a-1, N-a)$$

$$SST = \sum_i \sum_j (\bar{y}_{ij} - \bar{y}_{..})^2$$

$$SSE = \sum_i \sum_j (y_{ij} - \bar{y}_{ij})^2$$

ناحیه بحرانی :

$$R H_0 \rightarrow F_0 > F_{\alpha} (a-1, N-a)$$

مثال ۳-۱: صحنه ۵۷ کتاب : هدف : بررسی مقاومت کشش الیاف بر اساس (درصدینیه های) مختلف بکار رفته در الیاف

۱۵٪ / ۲۰٪ / ۲۵٪ / ۳۰٪ / ۳۵٪

واحد آزمایش : الیاف

عامل : یکنه عامل : مقاومت کشش که حاصل می شود

تیمار : (درصدینیه های بکار رفته در الیاف) (سطوح یکنه) ۵، ۲۰، ۲۵، ۳۰، ۳۵

تبعه شرایط یکسان فقط (درصدینیه ها در این آزمایش تغییر می کنند و برای هر کدام از تیمارها ۵ تکرار انجام گرفته است) . $n=5, 20, 25, 30, 35$

یک طرح تک عاملی با ۵ تکرار اجرا شده است . تعریف طرح تعاقبی ساده نوشته شود .

تعاقبی : $y_{ij} = \mu + T_i + \epsilon_{ij}$ $\epsilon_{ij} \sim N(0, \sigma^2)$ $\sum_{i=1}^5 T_i = 0$ مدل : $y_{ij} = \mu + T_i + \epsilon_{ij}$ $\epsilon_{ij} \sim N(0, \sigma^2)$ $\sum_{i=1}^5 T_i = 0$

$$\begin{cases} H_0 : T_1 = T_2 = \dots = T_5 = 0 \\ H_1 : \text{Not } H_0 \end{cases} \quad \begin{cases} H_0 : \mu_1 = \mu_2 = \dots = \mu_5 = \mu \\ H_1 : \text{Not } H_0 \end{cases}$$

تکرار	۱	۲	۳	۴	۵	جمع تیمار
۱۵٪	۷	۷	۱۵	۱۱	۹	۴۹
۲۰٪	۱۲	۱۷	۱۲	۱۸	۱۸	۷۷
۲۵٪	۱۴	۱۸	۱۸	۱۹	۱۹	۸۸
۳۰٪	۱۹	۲۵	۲۲	۱۹	۲۳	۱۰۸
۳۵٪	۱۵	۱۰	۱۱	۱۵	۱۱	۵۲
جمع تیمار						$\sum y_{ij} = 376$

$$N = a \times b = 5 \times 5 = 25$$

$$SST_0 = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{N} = 11^2 + 11^2 + \dots + 11^2 - \frac{(374)^2}{25} = 4292 - 5455.04 = 434.94$$

$$SST = \sum_{i=1}^a \frac{Y_i^2}{n} - \frac{Y_{..}^2}{N} = \frac{49^2 + 177^2 + 188^2 + 101^2 + 54^2}{5} - \frac{(374)^2}{25} = 4130.44 - 5455.04 = 475.17$$

$$SSE = SST_0 - SST = 434.94 - 475.17 = 141.2$$

تفسیر نتایج: با استفاده از آزمون F (همچنین روی مقادیر کسپی تأثیر می‌گذارند).
 خطا: 141.2 تیمار: 434.94 = 475.17 + 141.2 نتیجه: وابسته

جدول ANOVA طرح تعاملي ساده

Source	df	SS	MS	F
تیمار treatment	$a-1 = 5-1 = 4$	$SST = 475.17$	$MST = \frac{SST}{a-1} = \frac{475.17}{4} = 118.79$	$F_0 = \frac{MST}{MSE} = \frac{118.79}{11.04} = 10.75$
خطا Error	$N-a = 25-5 = 20$	$SSE = 141.2$	$MSE = \frac{SSE}{N-a} = \frac{141.2}{20} = 7.06$	
total	$N-1 = 24$	$SST_0 = 434.94$		

$$F_{(4,20)} = 2.187 \leftarrow RH_0 \leftarrow F_0 > F_{\alpha}(a-1, N)$$

نتیجه گیری: بر اساس شواهد موجود (طبق این داده‌ها یعنی داده غرض سوز حمله نتیجه غرض شود) با اطمینان 19.5 تفاوت معنی داری بین تیمارها وجود دارد.

نتیجه گیری کلی: بر اساس شواهد موجود با اطمینان 19.5 پهنه‌های مختلف روی مقاومت کسپی این الیاف تأثیرگذار است. اما اختلاف دارند؟ بهترین کدام است؟
 مقایسه زوجی

فرض کنید Z_i ها متغیرهای تصادفی مستقل باشند (برای اثبات اینکه SST و SSE مستقل اند) : فرض کنید Z_i ها متغیرهای تصادفی مستقل باشند

بطوریکه : $Z_i \sim N(0, 1)$ $\sum_{i=1}^p Z_i^2 = Q_1 + Q_2 + \dots + Q_p$ که در آن S_p و Q_i ها متغیرهای مستقل

با درجه آزادی p هستند (در آن صورت داریم) $\sum_{i=1}^p Z_i^2 = p$ $\Leftrightarrow$ Q_i ها مستقل اند و دارای توزیع کای دو با درجه آزادی p هستند

طبق آنالیز واریانس (ANOVA) داریم :

$$\sum Z_i^2 \Rightarrow SST = SST + SSE \Rightarrow \sum Z_i^2 = Q_1 + Q_p \Rightarrow p_1 + p_p = a-1 + N-a = N-1 \xrightarrow{\text{طبق قضیه کمران}} Q_i \text{ ها مستقل اند}$$

$p_1 = (N-1)$ $p_p = (a-1)$ $p_p = (N-a)$ SST و SSE مستقل اند

در طرح تصادفی ساده (متداول) :

$$E(MST) = \sigma^2 + n \frac{\sum_{i=1}^a T_i^2}{a-1}$$

اثبات :

$$E(MSE) = \sigma^2$$

$$E(MSE) = E\left(\frac{SSE}{N-a}\right) = \frac{1}{N-a} E(SSE) = \frac{1}{N-a} E\left(\sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \sum_{i=1}^a \frac{y_{i.}^2}{n}\right) = \frac{1}{N-a} \left[\sum_{i=1}^a \sum_{j=1}^n E(y_{ij}^2) - \sum_{i=1}^a \frac{E(y_{i.}^2)}{n} \right] =$$

$$\frac{1}{N-a} \left[\sum_{i=1}^a \sum_{j=1}^n (\mu_i^2 + T_i^2 + \sigma^2 + 2\mu_i T_i) - \sum_{i=1}^a \frac{(n\mu_i^2 + nT_i^2 + n\sigma^2 + 2n\mu_i T_i)}{n} \right] = \frac{1}{N-a} \left[an\mu^2 + n \sum_{i=1}^a T_i^2 + an\sigma^2 + 2n\mu \sum_{i=1}^a T_i - an\mu^2 - n \sum_{i=1}^a T_i^2 - an\sigma^2 - 2n\mu \sum_{i=1}^a T_i \right] =$$

$$\frac{1}{N-a} (an\sigma^2 - an\sigma^2) = \frac{1}{N-a} (an\sigma^2 - an\sigma^2) = \sigma^2$$

از جمله بندی

$$E(y_{ij}^2) = E((\mu + T_i + \epsilon_{ij})^2) = E(\mu^2 + T_i^2 + \epsilon_{ij}^2 + 2\mu T_i + 2\mu \epsilon_{ij} + 2T_i \epsilon_{ij}) = E(\mu^2) + E(T_i^2) + E(\epsilon_{ij}^2) +$$

$$E(2\mu T_i) + E(2\mu \epsilon_{ij}) + E(2T_i \epsilon_{ij}) = \mu^2 + T_i^2 + \sigma^2 + 2\mu T_i + 2\mu E(\epsilon_{ij}) + 2T_i E(\epsilon_{ij}) = \mu^2 + T_i^2 + \sigma^2 + 2\mu T_i$$

$$y_{i.} = \sum_{j=1}^n y_{ij} \quad \left. \begin{array}{l} E(\epsilon_{ij}^2) = \sigma^2 + 0 = \sigma^2 \\ E(\epsilon_{ij}) = \text{var}(\epsilon_{ij}) + [E(\epsilon_{ij})]^2 \end{array} \right\} E(x^2) = \text{var}(x) + [E(x)]^2$$

P4PCO

$$\text{var}(x_1 + x_p) = \text{var}(x_1) + \text{var}(x_p) + 2\text{cov}(x_1, x_p) \Rightarrow \text{var}(\epsilon_{ij}) = n\sigma^2$$

توزیع t دارد به آزادی کوفاک $\rightarrow$ توزیع فنیتر دارد به آزادی کوفاک $\rightarrow$ توزیع گوسی $\rightarrow$ قضیه حد مرکزی برای این ها برقرار نیست

Subject:

Date:

$$E(\sum \dots) = \sum E(\dots) \leftarrow \text{اگر در این خطی بودن دارد یعنی}$$

اگر اید یک متغیر تصادفی باشد، توان (دوم) اید با خطی آن = واریانس آن متغیر است

$$\langle 143 / 7 / 22 \rangle$$

$$y_{ij}^* = a y_{ij} + b \quad \text{یا} \quad y_{ij}^* = \frac{y_{ij}}{a} - b \quad a, b \in \mathbb{R} \quad \text{داده های که ندارند شده}$$

کم و زیاد کردن داده ها (+ کردن) تاثیر روی SS ها ندارد پس روی F هم تاثیر ندارد

$$F_o^* = \frac{\frac{SS}{a^2}}{\frac{SS}{a^2}} = F_o$$

اگر داده ها را ضرب و تقسیم کنیم F_o^* با F_o اولیه تفاوتی ندارد

$$SS^* = a^2 SS$$

ضرب کنیم $SS \leftarrow$ جدید $\frac{1}{a}$ برابر SS قدیم می شود

$$SS = \frac{SS^*}{a^2}$$

تقسیم کنیم $SS \leftarrow$ جدید $\frac{1}{a}$ برابر SS قدیم می شود

$$Z_{ij} = \frac{y_{ij} - y_{..}}{S_{ij}}$$

داده های استاندارد شده

در کل هر نوع تبدیل خطی روی داده ها تاثیری روی F نمی گذارد

اثبات MST:

$$y_{ij} = \mu + T_i + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

$$\varepsilon_{ij} \text{ iid } N(0, \sigma^2)$$

$$\sum_{i=1}^n T_i = 0$$

$$E(MST) = \sigma^2 + n \frac{\sum_{i=1}^n T_i^2}{a-1}$$

$$SST_f = \sum_{i=1}^a \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{an}$$

$$E(MST) = E\left(\frac{SST}{a-1}\right) = \frac{1}{a-1} E(SST) = E\left(\frac{\sum_{i=1}^a \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{an}\right) = \frac{\frac{1}{n} \sum_{i=1}^a E(y_{i.}^2) - \frac{1}{an} E(y_{..}^2)}{a-1} =$$

$$\frac{1}{a-1} \left[\frac{1}{n} \sum_{i=1}^a (n^2 \mu^2 + n^2 T_i^2 + n \sigma^2 + 2n \mu T_i) - \frac{(an^2 \mu^2 + an \sigma^2)}{an} \right] = \frac{1}{a-1} \left[(an \mu^2 + n \sum_{i=1}^a T_i^2 + an \sigma^2 + 0) - an \mu^2 \right]$$

$$PAPCO \quad \frac{n \sum_{i=1}^a T_i^2}{a-1} + \frac{\sigma^2 (a-1)}{(a-1)} = \sigma^2 + n \frac{\sum_{i=1}^a T_i^2}{a-1}$$

$$E(y_{ij}) \stackrel{\text{حداقل}}{=} n^2 \mu^2 + n^2 T_i^2 + n^2 \sigma^2 + 2n^2 \mu T_i$$

$$E(y_{..}) = E((an\mu + \epsilon_{..})) = E(an^2\mu^2 + 2an\mu\epsilon_{..} + \epsilon_{..}^2) = an^2\mu^2 + 2an\mu E(\epsilon_{..}) + E(\epsilon_{..}^2) =$$

$$an^2\mu^2 + an\sigma^2 \quad y_{..} = \sum_i \sum_j y_{ij} = \sum_{i=1}^a \sum_{j=1}^n (\mu + T_i + \epsilon_{ij}) = an\mu + n \sum T_i + \epsilon_{..}$$

$$\text{var}(\epsilon_{..}) = \text{var}\left(\sum_{i=1}^a \sum_{j=1}^n \epsilon_{ij}\right) = \sum_{i=1}^a \sum_{j=1}^n \text{var}(\epsilon_{ij}) + \sum_{i \neq i'} \sum_{j \neq j'} \sum_j \sum_{j'} \text{cov}(\epsilon_{ij}, \epsilon_{i'j'}) = an\sigma^2$$

برای درک بهتر اینها، مدل طرح متعادل (بالا) را با اثرات تصادفی مقایسه کنید.

$$y_{ij} = \mu + T_i + \epsilon_{ij}$$

$$\epsilon_{ij} \sim N(0, \sigma^2) \text{ شرطی}$$

$$\sum_{i=1}^a T_i = 0$$

$$\mu \Rightarrow \text{میانگین} \quad T_i \Rightarrow \text{تفاوت} \quad \sigma^2 \Rightarrow \text{پراش}$$

$$\hat{\sigma}^2 = \text{MSE} \quad E(\text{MSE}) = \sigma^2$$

روش حداقل مربعات (برای خطاها)

$$\min_{\mu, T_i} \sum_i \sum_j \epsilon_{ij}^2 = \sum_i \sum_j (y_{ij} - \mu - T_i)^2$$

$$\frac{\partial f}{\partial \mu} = -2 \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{T}_i) = 0 \Rightarrow y_{..} - an\hat{\mu} - n \sum_{i=1}^a \hat{T}_i = 0 \Rightarrow \hat{\mu} = \frac{y_{..}}{an} = \bar{y}_{..}$$

$$\frac{\partial f}{\partial T_i} = -2 \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{T}_i) = 0 \Rightarrow y_{i.} - n\hat{\mu} - n\hat{T}_i = 0 \Rightarrow \hat{T}_i = \frac{y_{i.} - n\bar{y}_{..}}{n} = \bar{y}_{i.} - \bar{y}_{..}$$

$$\hat{y}_{ij} = \hat{\mu} + \hat{T}_i = \bar{y}_{..} + \bar{y}_{i.} - \bar{y}_{..} = \bar{y}_{i.}$$

$$e_i = y_{ij} - \hat{y}_{ij}$$

PAPCO

باقی‌مانده‌ها که حاصل شده‌اند، شرایط مدل را داشته‌اند.

چک شود که آیا باقی‌مانده‌ها

- ۱- نرمالند
- ۲- هم‌وابسته‌اند
- ۳- ناهمبسته‌اند

با نرم افزار مینی‌تب چک می‌کنیم.

بررسی شرایط مدل (مناسبت مدل) ۵ در طرح تصادفی ساده $\leftarrow$
 $y_{ij} = \mu + T_i + \epsilon_{ij}$

۳ شرط اساسی که باید روی خطاها برقرار باشد $\leftarrow$

- ۱- نرمال بودن خطاها
- ۲- هم واریانس خطاها
- ۳- ناهمبسته بودن

$\epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$

چرا باید شرایط مدل بررسی شود؟

۱) اگر خطاها نرمال نباشند چه اتفاقی می افتد؟ توزیع آماره فیشر دچار مشکل می شود. آزمون فرض ها اعتبار ندارند.

۲) اگر واریانس خطاها یکسان نباشد چه اتفاقی می افتد؟
 $F = \frac{MST}{MSE} \stackrel{?}{\sim} \frac{SST}{\sigma^2} \sim \chi^2_{(a-1)}$
 $E(MSE) = \sigma^2$
 $E(MST) = ?$
توزیع F پیچیده می شود.

۳) ناهمبسته بودن خطاها اگر برقرار نباشد چه اتفاقی می افتد؟ توزیع آماره آزمون فیشر مشخص نمی شود. $F \sim ?$

* در همه این اشیاء ها از ناهمبسته بودن استفاده کرده ایم *

توجه: حد درستی ۵ اگر تعداد نمونه ها زیاد باشد، توزیع میانگین جامعه نرمال می شود.

چندک نرمال با چندک داده ها یکی باشد می گوئیم نرمال است $\leftarrow$ $q-q \text{ plot}$

چندک ها مهم تر از شکل توزیع است $\leftarrow$ مثلاً $Beta(2,2)$ شکل آن نرمال نیست اما شکل توزیع شبیه نرمال است.

$q-q \text{ plot}$ میار چندک داده ها را در مقابل چندک نرمال رسم می کنند یعنی مثلاً چندک ۰.۰۵ داده داده را در مقابل چندک ۰.۰۵ نرمال رسم می کنیم.

در آزمون لکونوف-شیمیسه باید میانگین و واریانس بدویم.

PAPCO

هیچ $= 0$ است. یعنی باقی مانده ها مشکوک هستند.
خطای استاندارد کمتر از ۳ باشد مشکل دارد یا بدین ۲ و ۲ باشد.
چون نرمال پس ۳ و ۳ نیست.

$$e_{ij}^* = \frac{e_{ij} - \bar{e}}{\sqrt{MSE}} = \frac{e_{ij}}{\sqrt{MSE}}$$

ناصبه بودن $\leftarrow$ استقلال $\rightarrow$ تحت شرایط خاص و مثلا توزیع نرمال

استقلال $\leftarrow$ نتیجه دهد $\leftarrow$ ناصبه بودن

کلی تر است

Subject:
Date:

راه کارها:

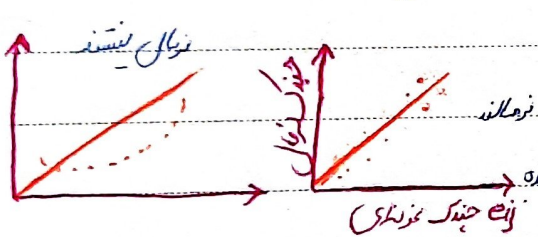
① اگر خطاها نرمال شود $\leftarrow$ $\left. \begin{array}{l} 1- \text{تبدیل نرمال} \text{ (برای نمونه های زیاد)} \text{ } \rightarrow \text{ } \text{box-cox} \\ 2- \text{تایدار هستی} \text{ (برای مشاهدات کم)} \end{array} \right\}$

② اگر خطاها هم واریانس نباشند $\leftarrow$ $\left. \begin{array}{l} 1- \text{تبدیل هم واریانس} \text{ } \rightarrow \text{ } \text{box-cox} \\ 2- \text{روش هایی که ناهمبندی واریانس را حل کنند} \end{array} \right\}$

③ اگر خطاها ناصبه بودن $\leftarrow$ باید به سراغ مدل های جایگزین برویم که همبستگی را مدل کنند.

پیرامونی مناسب مدل در طرح تصادفی ساده:

$\leftarrow$ باقیمانده های اصل



$$\bar{y}_1 - \bar{y}_2 = \bar{y}_1 - \bar{y}_2 = \bar{y}_1 - \bar{y}_2$$

1- باقیمانده های مدل را جایگزین کنید.

1- روش نموداری $\leftarrow$ $a \cdot plot$ باقیمانده $\rightarrow$ e چندک نموداری

1- جلب کردن نرمالیتی باقیمانده ها

چندک آن حول خط میانه باشد تا نرمال باشد

چندک آن نزدیک 0 باشد تا نرمال باشد

داده های بزرگ یا داده های پرت روی نرمال بودن تأثیر می گذارند باید بررسی کنیم

1- شاپیرو ویلیک $\rightarrow$ پارامتری

2- لکونوف $\rightarrow$ ناپارامتری

3- آندرسن دارلینگ AD

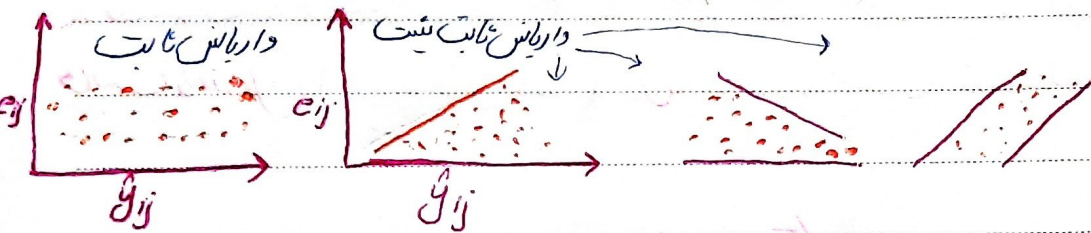
در R باید یکجای نصب کنیم

فرض H در این آزمون ها $\rightarrow$ باقی مانده ها نرمال هستند

طرح رابانگ انفرار Mini Tab (هم برای دوره وکالت)

۲- هم‌دارایی خطا (ثابت بودن) واریانس خطاها

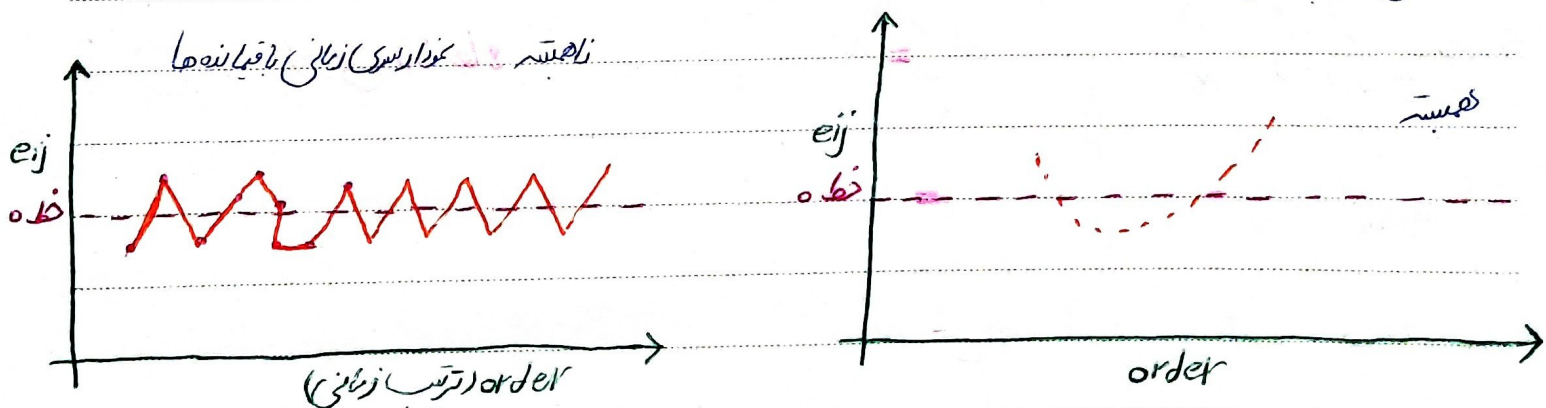
- چک کردن واریانس ثابت خطاها (باقی‌مانده‌ها)
- ۱- روش نموداری: اگر هیچ الگوی مشاهده نگردیم یعنی واریانس ثابت است.
 اگر الگوی قیفی شکل مشاهده کردیم، واریانس ثابت نیست.



۲- روش عددی (آزمون) a-تایار $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_a^2$

- $a=2$ ← آزمون فشرانجامی (هم)
- $a=2$ ← آزمون بارتلیت (بارامتری) خطا نهال بود از بارتلیت می‌رویم
- $a=2$ ← آزمون بولین (نا بارامتری) خطا نهال نبود از این راه می‌رویم

۳- بررسی ناهمبسته بودن باقی‌مانده‌ها: نموداری



- باید به طور کامل تقاضای حول صفر برانگیز شده باشند.
- هرگونه الگوی مشاهده شود، یعنی همبسته هستند.

طرح متعالی بهتر است یا نامتعالی؟ طرح متعالی

مزیت های طرح متعادل؟ ① توان آزمون بیشتر است نسبت به طرح نامتعادل (خطای نوع دوم - ۱ - توان آزمون)

② انحرافات جزئی از نا هم دارایی خطاها در طرح متعادل قابل چشم پوشی است.

« ۱۳۵۳ / ۸ / ۵ »

حالت نامتعادل (طرح تصادفی ساده) مدل با اثرات ثابت شده و اگر تعداد تکرارها به هر دلیل برای تیمارها

متفاوت شوند طرح نامتعادل است. تکرار تیمار i ام n_i تعداد کل مشاهدات $N = \sum_{i=1}^a n_i$

(حسابی کمتر)

طرح متعادل دو مزیت نسبت به طرح نامتعادل دارد { ۱ - نسبت به انحرافات جزئی از برابری دارایی ها استوار است. ۲ - آزمون ها توانمندتر است. }

مدل، شرایط مدل و ... کلی یکسان است یعنی مشابه طرح متعادل است. فقط هر جا n است تبدیل به n_i می شود

$$\sum_{i=1}^a n_i T_i = 0 \quad \text{شرایط مدل}$$

$$SST_0 = \sum_{i=1}^a \sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N} \quad df_{T_0} = N - 1 \quad \text{فرمول های محاسباتی SS ها}$$

$$SST = \sum_{i=1}^a \frac{y_i^2}{n_i} - \frac{y_{..}^2}{N} \quad df_T = a - 1$$

$$SSE = SST_0 - SST \quad df_E = N - a$$

M.S ها قبل است و فرمول آن تغییر نمی کند.

$$E(MSE) = \sigma^2$$

$$E(MST) = \sigma^2 + \frac{\sum_{i=1}^a n_i T_i^2}{a - 1}$$

SPSS با Post hoc مقایسات زوجی را انجام می دهد. LSD همان t تست است

Subject:
Date:

مقایسات زوجی میانگین تبارها (آزمون تعقیبی) *

آزمون مقایسه ها $\rightarrow$ ضریب ترتیب ها برابر باشد. آزمون مقایسه را با C نشان می دهیم. اگر آزمون های که

ایجاد می دهیم آزمون مقایسه است. $\mu_1 = \mu_2 \Rightarrow \mu_1 - \mu_2 = 0$

ضریب که آن ها است $\mu_1 + \mu_2 = \mu_3 + \mu_4 \Rightarrow \mu_1 + \mu_2 - \mu_3 - \mu_4 = 0$

$$H_0: \sum_{i=1}^k C_i \mu_i = 0$$

$$\sum_{i=1}^k C_i = 0$$

$$\begin{cases} H_0: \mu_i = \mu'_i \\ H_1: \mu_i \neq \mu'_i \end{cases}$$

آزمون تعقیبی $\leftarrow$

چه وقت برای مقایسات زوجی (آزمون های تعقیبی) می رویم؟ زمانی که فرض H_0 در ANOVA معنی دار نشود (لاستور)

یعنی تفاوت معنی دار بین تبارها وجود دارد

تعداد تبارها $\begin{pmatrix} a \\ b \end{pmatrix}$ آزمون داریم. چند مقایسه زوجی داریم *

- 1- LSD Least square Differents (یعنی حداقل تفاوت معنی دار)
 - 2- دانکن
 - 3- نیوون - کوک
 - 4- توکی
 - 5- دانست همه تبارها با یک تبار مشاهده مقایسه می شوند
- چه آزمون هایی وجود دارد $\leftarrow$

تبار مشاهده (تبار کنترل) * یعنی چیزی که روی آن چیز جدید اعمال نمی شود.

$$H_0: \mu_i = \mu_j$$

روش LSD

$$|t_o| > t_{\frac{\alpha}{2}}(df_E) \quad t_o = \frac{\bar{Y}_i - \bar{Y}_j}{\sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}$$

آماره آزمون

$$|\bar{Y}_i - \bar{Y}_j| > LSD = t_{\frac{\alpha}{2}}(df_E) \times \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \Rightarrow \text{تفاوت معنی داری بین تیمارها وجود دارد}$$

تعیین معنی داری

مثال: داده های مثال ۱-۳. در این مثال فرض H_0 محذوف. ANOVA روش. $\alpha = 0.05$. $n_i = n = 5$

$$df_E = N - a = 20 \quad \bar{a} = 5 \quad MSE = 110.4 \quad \left(\frac{5}{2} \right) = 10$$

$$\bar{Y}_1 = 9.8 \quad \bar{Y}_2 = 15.4 \quad |\bar{Y}_1 - \bar{Y}_2| = |9.8 - 15.4| = 5.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$\bar{Y}_1 = 9.8 \quad \bar{Y}_3 = 17.4$$

$$|\bar{Y}_1 - \bar{Y}_3| = |9.8 - 17.4| = 7.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$\bar{Y}_2 = 15.4 \quad \bar{Y}_4 = 21.4$$

$$|\bar{Y}_2 - \bar{Y}_4| = |15.4 - 21.4| = 6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$\bar{Y}_3 = 17.4 \quad \bar{Y}_5 = 10.8$$

$$|\bar{Y}_1 - \bar{Y}_5| = |9.8 - 10.8| = 1 < LSD = 3.75 \quad \text{معنی دار نیست}$$

$$\bar{Y}_4 = 21.4 \quad \bar{Y}_5 = 10.8$$

$$|\bar{Y}_2 - \bar{Y}_5| = |15.4 - 10.8| = 4.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$|\bar{Y}_3 - \bar{Y}_5| = |17.4 - 10.8| = 6.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$|\bar{Y}_4 - \bar{Y}_5| = |21.4 - 10.8| = 10.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$|\bar{Y}_1 - \bar{Y}_4| = |9.8 - 21.4| = 11.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$|\bar{Y}_2 - \bar{Y}_5| = |15.4 - 10.8| = 4.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$|\bar{Y}_3 - \bar{Y}_5| = |17.4 - 10.8| = 6.6 > LSD = 3.75 \quad * \text{ معنی دار است}$$

$$LSD = t_{\frac{\alpha}{p}}(df_E) \times \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

$$LSD = t_{\frac{\alpha}{p}}(df_E) \times \sqrt{MSE \cdot 2n} \leftarrow \text{طرح متعادل} \quad LSD = 2.1084 \times \sqrt{1.104 \times \left(\frac{4}{5} \right)} = 3.175$$

$$t_{0.05/25}(25) = 2.1084$$

$$\text{ترتیب} = \bar{Y}_1, \bar{Y}_5, \bar{Y}_2, \bar{Y}_3, \bar{Y}_4$$

* حجم گروه بندی ۵، کوچکترین میانگین را در گروه اول قرار می دهیم.

گروه اول	گروه دوم	گروه سوم
$\bar{Y}_1$	$\bar{Y}_2$	$\bar{Y}_4$
$\bar{Y}_5$	$\bar{Y}_3$	

یعنی با فاصله اختلاف معنی دارد

چون اختلاف معنی دارد نمی تواند اینی قرار گیرد

بین گروه ها اختلاف معنی داری وجود دارد
در هر گروه ها اختلاف معنی داری وجود ندارد

کدام تیم بهتر است؟ انتخاب ملاک بهترین مشخص شود که نزدیکتر به صفر باشد

در این مثال هر چه مقاومت کمتری بیشتر باشد بهتر است پس هر چه بیشتر باشد بهتر است پس گروه سوم بهتر است چون

مقاومتش از بقیه بیشتر است. یعنی بنی ۲. در حد بهتر از بقیه است.

روش دانکن:

۱- میانگین تیمها را از کوچک به بزرگ مرتب می کنیم ۲- تعیین مرتبه ی مقایسه دو میانگین ها ۳- $ST_{ij} = \sqrt{\frac{MSE}{n}}$

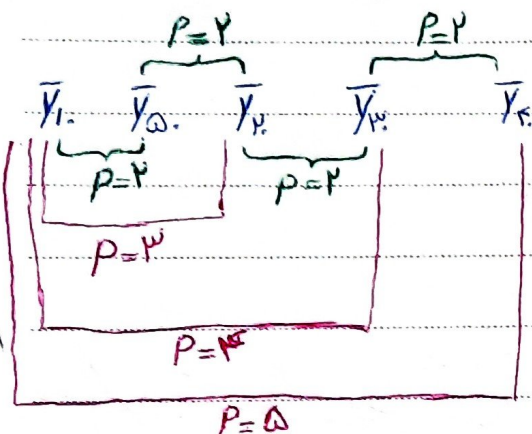
$$Rp = t_{\alpha}(p, df_E) \cdot ST_{ij}$$

$$n = \frac{a}{\sum_{i=1}^a \frac{1}{n_i}}$$

انجمن دانش آفرینان

اگر $R_p > |\bar{y}_i - \bar{y}_j|$ اختلاف معنی داری دارند.

تست p به ترتیب آمارهای زیر را می نویسیم. سپس از نتایج که کنار هم هستند $p=2$ هست.



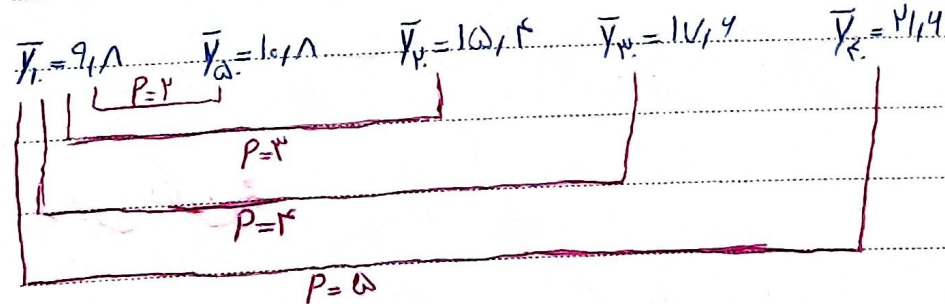
۱۴۰۳/۸/۶

ادامه رویش دانلش.

۵- قدر مطلق اختلاف میانگین های تیماری را با R_p مقایسه می کنیم. اگر بزرگتر بود، معنی دار هستند. $R_p > |\bar{y}_i - \bar{y}_j|$

۴- گروه بندی می کنیم معتبرترین گروه را انتخاب می کنیم.

مثال دانلش: ۵ ادامه مثال ۱-۳ در رویش ها



$$a = 5 \quad n = 5$$

$$MSE = 110.4$$

$$df_E = 25 - 5 = 20$$

$$n = \frac{a}{\sum_{i=1}^a \frac{1}{n_i}} = \frac{5}{5 \times \frac{1}{5}} = 5$$

$$S\bar{y}_i = \sqrt{\frac{MSE}{n}} = \sqrt{\frac{110.4}{5}} = 14.7$$

$$R_p = t_{\alpha}(p, df_E) \times S \bar{y}_i \Rightarrow R^p = t_{\alpha}(p, 20) \times 1.127 = 2.95 \times 1.127 = 3.325$$

از جدول دانستن آمار

* همیشه در طرح متقابل $R_p = LSD$ است *

$$R_p = t_{\alpha}(p, 20) \times S \bar{y}_i = 3.10 \times 1.127 = 3.494$$

* هر چه p افزایش می‌دهیم باید R_p افزایش می‌دهیم پس R_p است $p=0.05$

$$R_p = t_{\alpha}(p, 20) \times 1.127 = 3.18 \times 1.127 = 3.584$$

* صغری است نتیجه: $R_p > LSD$ *

$$R_p = t_{\alpha}(p, 20) \times 1.127 = 3.25 \times 1.127 = 3.664$$

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 5.12 > R_p = 3.664$$

* معنی دار است *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 7.18 > R_p = 3.664$$

* معنی دار است *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 11.18 > R_p = 3.664$$

* معنی دار است *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 1 < R_p = 3.664$$

* معنی دار نیست *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 2.12 < R_p = 3.664$$

* معنی دار نیست *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 4.12 > R_p = 3.664$$

* معنی دار است *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 4.14 > R_p = 3.664$$

* معنی دار است *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 4 > R_p = 3.664$$

* معنی دار است *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 4.18 > R_p = 3.664$$

* معنی دار است *

$$p=0.05 \Leftarrow |\bar{y}_1 - \bar{y}_p| = 10.18 > R_p = 3.664$$

* معنی دار است *

گروه اول	گروه دوم	گروه سوم	بهترین گروه کدام است؟ هر چه تفاوت بیشتری با بقیه داشته باشد
$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	
$\bar{y}_5$	$\bar{y}_4$		← هر چه میانگین بیشتری داشته باشد. گروه سوم تبار از قبیل بهتر است.

* همیشه نتیجه دانستن و LSD یکی نمی شود و ممکن است متفاوت باشند *

روشن نویسی - کونتر: مشابه روشن دانستن است. (همه موارد مثل روشن دانستن)

$$k_p = q(p, df_E) \cdot S\bar{y}_i$$

$$p=2 \Rightarrow k_p = R_p = LSD$$

$$p>2 \Rightarrow k_p > R_p > LSD \quad \text{ارتباط با دانستن و LSD}$$

داده های استوار نیست (مثله)

$$S\bar{y}_i = 1.127$$

$$df_E = 20$$

$$\alpha = 0.105$$

مثال روشن نویسی - کونتر: مثال ۳-۱

۱- مشابه روشن دانستن مثل می کنیم ۱:۰ - ۲- $S\bar{y}_i = \sqrt{\frac{MSE}{n}}$ - ۳- محاسبه k_p

$$k_p = q(2, 20) \times 1.127 = R_p = LSD = 2.175$$

$$k_p = q(3, 20) \times 1.127 = 2.58 \times 1.127 = 2.905$$

$$k_p = q(4, 20) \times 1.127 = 3.94 \times 1.127 = 4.44$$

$$k_p = q(5, 20) \times 1.127 = 4.22 \times 1.127 = 4.75$$

$$p=2 \quad |\bar{y}_1 - \bar{y}_2| = 5.14 > k_p = 2.905$$

* معنی دار است

$$p=4 \quad |\bar{y}_1 - \bar{y}_4| = 7.18 > k_p = 4.44$$

* معنی دار است

$$P=5 \quad |\bar{Y}_1 - \bar{Y}_5| = 11.8 > K_5 = 5.38 \quad * \text{ معنی دار است}$$

$$P=2 \quad |\bar{Y}_1 - \bar{Y}_5| = 1 < K_2 = 3.75 \quad \text{معنی دار نیست}$$

$$P=2 \quad |\bar{Y}_1 - \bar{Y}_3| = 2.2 < K_2 = 3.75 \quad \text{معنی دار نیست}$$

$$P=3 \quad |\bar{Y}_1 - \bar{Y}_3| = 4.2 > K_3 = 4.55 \quad * \text{ معنی دار است}$$

$$P=2 \quad |\bar{Y}_1 - \bar{Y}_5| = 4.4 > K_2 = 3.75 \quad * \text{ معنی دار است}$$

$$P=2 \quad |\bar{Y}_1 - \bar{Y}_4| = 4 > K_2 = 3.75 \quad * \text{ معنی دار است}$$

$$P=3 \quad |\bar{Y}_1 - \bar{Y}_5| = 4.8 > K_3 = 4.55 \quad * \text{ معنی دار است}$$

$$P=4 \quad |\bar{Y}_1 - \bar{Y}_5| = 4.8 > K_4 = 5.02 \quad * \text{ معنی دار است}$$

گروه اول	گروه دوم	گروه سوم
$\bar{Y}_1$	$\bar{Y}_2$	$\bar{Y}_3$
$\bar{Y}_4$	$\bar{Y}_5$	

گروه بندی:

روش توکی (Tukey): مشابه روش فیشر، گوناگون است. فقط آماره تفاوت و فقط یک مرتبه در نظر

گرفته می شود. این روش نسبت به دانکن و فیشر گوناگون تر است و گوناگون تر است.

$$p=a \Rightarrow T_\alpha = K_\alpha$$

مقایسه با بقیه:

$$p=2 \Rightarrow T_\alpha > K_p = R_p = LSD$$

$$p>2 \Rightarrow T_\alpha > K_p > R_p > LSD$$

برای محاسبه ضرایب همبستگی باید مقدار داریم.

$$T_{\alpha} = q_{\alpha}(a, df_E) \cdot S_{\bar{Y}_i}$$

در صفتی تب تستی فرض از روش تولی استفاده می شود.

مثال تولی ۸: مثال ۱-۳. مراحل محاسباتی همانند جدول - کوتاه است - فقط مرتبه نیاز داریم.

$$K_{\alpha} = T_{\alpha} = q_{\alpha}(5, 20) \times S_{\bar{Y}_i} = 4, 24 \times 1, 27 = 5, 38$$

$|\bar{Y}_1 - \bar{Y}_2| = 5, 4 > T_{\alpha} = 5, 38$ * معنی دار است

$|\bar{Y}_1 - \bar{Y}_3| = 7, 8 > T_{\alpha} = 5, 38$ * معنی دار است

$|\bar{Y}_1 - \bar{Y}_4| = 11, 8 > T_{\alpha} = 5, 38$ * معنی دار است

$|\bar{Y}_1 - \bar{Y}_5| = 1 < T_{\alpha} = 5, 38$ معنی دار نیست

$|\bar{Y}_2 - \bar{Y}_3| = 2, 2 < T_{\alpha} = 5, 38$ معنی دار نیست

$|\bar{Y}_2 - \bar{Y}_4| = 6, 2 > T_{\alpha} = 5, 38$ * معنی دار است

$|\bar{Y}_2 - \bar{Y}_5| = 4, 4 < T_{\alpha} = 5, 38$ معنی دار نیست

$|\bar{Y}_3 - \bar{Y}_4| = 4 < T_{\alpha} = 5, 38$ معنی دار نیست

$|\bar{Y}_3 - \bar{Y}_5| = 4, 8 > T_{\alpha} = 5, 38$ * معنی دار است

$|\bar{Y}_4 - \bar{Y}_5| = 10, 8 > T_{\alpha} = 5, 38$ * معنی دار است

لذا اگر گروه معتبر است ؟ هر دو مطابق نیست، معتبر

← گروه چهارم ← تیمار ۳ و تیمار ۴ معتبر است

گروه اول	گروه دوم	گروه سوم	گروه چهارم
$\bar{y}_1$	$\bar{y}_5$	$\bar{y}_2$	$\bar{y}_3$
$\bar{y}_5$	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$

توکی نیست به یاد کردن فرض H_0 نسبت به بقیه می‌توان کار تر است.

روش دانست : همه تیمارها با تیمار شاهد مقایسه می‌شود. فرض می‌کنیم تیمار ۱ را شاهد باشد.

$$\begin{cases} H_0 : \mu_i = \mu_a \\ H_1 : \mu_i \neq \mu_a \end{cases}$$

$$|\bar{y}_i - \bar{y}_a| > d_{\alpha}(a-1, df_E) \times \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_a} \right)} \leftarrow \text{اختلاف معنی دار است} \quad \text{تصمیم گیری}$$

از جدول دانست محاسبه می‌شود.

۱۴۰۳ / ۸ / ۱۲

$$\begin{cases} H_0 : \mu_i = \mu_5 \\ H_1 : \mu_i \neq \mu_5 \end{cases} \quad \text{مثال روش دانست : فرض کنید در مثال ۱-۳ تیمار پنجم تیمار شاهد باشد.}$$

در این طرح چون مقادیر معنی آماری دانست برای همه یکسان است.

$$\alpha = 0.05, \alpha = 5, df_E = N - a = 20, MSE = 1104, n = 5$$

$$d_{\alpha}(a-1, df_E) \cdot \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_a} \right)}$$

$$d_{0.05}(4, 20) \cdot \sqrt{MSE \times \frac{2}{n}} = 2.45 \times \sqrt{1104 \times \frac{2}{5}} = 41.79$$

PAPCO

از جدول دانست

$$|\bar{Y}_1 - \bar{Y}_5| = |918 - 1018| = 1 < 4.74 \quad \text{معنی دار نیست}$$

$$|\bar{Y}_2 - \bar{Y}_5| = |1514 - 1018| = 4.4 < 4.74 \quad \text{معنی دار نیست}$$

$$|\bar{Y}_3 - \bar{Y}_5| = |1744 - 1018| = 7.2 > 4.74 \quad * \text{معنی دار است}$$

$$|\bar{Y}_4 - \bar{Y}_5| = |2112 - 1018| = 10.9 > 4.74 \quad * \text{معنی دار است}$$

نتیجه: بین شاهد (تیمار ۵) با اطمینان ۹۵٪ بر اساس شواهد موجود با تیمارهای (سوک و چهارم) اختلاف معنی دار دارد اما با تیمارهای (۱ و ۲) اختلاف معنی داری ندارد.

مثال ۳:
 $T_1: (-1.0) \text{ یا } (0.5 \text{ و } 0.5) \text{ یا } (2.0 -)$
آلتر در بازه فاصله اطمینان باشد یعنی ممکن است تیمار ۱ اثر معنی دار نداشته باشد.

آلتر پایین (۰.۵ و ۰.۵-) باشد یعنی تیمار ۱ با اثر مثبت است.

آلتر بالا (۰.۵ - ۲.۰-) باشد یعنی تیمار ۱ با اثر منفی است.

آلتر در بازه بود یعنی تفاوت معنی داری ندارند. آلتر در بازه نبود یعنی تفاوت معنی دار دارند.

طرح ۱

- ۱- مدل با اثرات ثابت شده $\leftarrow$ اثر T_i ها ثابت هستند در مدل با اثرات ثابت شده
 - ۲- مدل با اثرات تصادفی شده $\leftarrow$ T_i ها تصادفی هستند. یعنی تیارها به تصادف از یک جامعه آماری انتخاب شده باشند.
 - ۳- مدل آمیخته $\leftarrow$ در مدل های سیرازیک عامل تغییر می شود. یعنی عامل ها تصادفی و بعضی دیگر ثابت شده هستند.
- طرح ۲

مدل طرح تصادفی ساده :

$$y_{ij} = \mu + T_i + \epsilon_{ij}$$

(متغیرهای کل را حذف کنید) $\rightarrow$ خطا تصادفی

T_i تعیین می کند که مدل با اثرات ثابت شده است یا تصادفی

تیارها به تصادف روی واحد های آزمایشی اعمال می شوند و این ربطی به مدل با اثرات تصادفی نشده ندارد.

نکته: تصادفی کردن در مسئله آزمایشی (حتما باید به انتخاب تیارها توجه کرد) که بگوییم مدل با اثرات تصادفی شده است.

مدل با اثرات تصادفی (طرح تصادفی ساده) :

مدل طرح :

$$y_{ij} = \mu + T_i + \epsilon_{ij}$$

شرط مدل :

$$N(0, \sigma^2) \text{ و } \epsilon_{ij}$$

در مدل با اثرات تصادفی $\sum T_i = 0$ نیست و $\leftarrow$ $N(0, \sigma^2)$ T_i

ناهمبستگی $\text{COV}(\epsilon_{ij}, T_i) = 0$

فرضیات (آزمون فرض) :

چون بداند تصادفی است صادی :

یعنی بولین با واریانس متناهی می باشد :

$$\begin{cases} H_0 : \sigma^2_T = 0 \\ H_1 : \sigma^2_T > 0 \end{cases}$$

P4PCO

* CoV (متغیر تصادفی و عدد ثابت) = 0 *

Subject:

Date:

متغیر تصادفی ثابت

اقتضای اینکه اول متغیر تصادفی ثابت باشد. $p(T_i = 0) = 1 \xrightarrow{\text{نتیجه}} p(X = E(X)) = 1$ است. $p(T_i = 0) = 1 \xrightarrow{\text{نتیجه}} p(X = E(X)) = 1$

آزاد واریانس: قبول SS ها و درجه آزادی قبل (میل با اثرات ثابت شده) می باشد. F_0 و تابع بحرانی

هم قبل است. $SST_0 = SST + SSE$

$F_0 = \frac{MST}{MSE} \sim F(a-1, n-a)$

$R.H. \Rightarrow F_0 > F_{\alpha}(a-1, n-a)$

امید MS های متفاوت است 8

$E(MST) = \sigma^2 + n\sigma_T^2$

$\sigma^2 = MSE$

برآورد ناآرایی

$E(MSE) = \sigma^2$

تحت فرض H_0 : $E(MST) = E(MSE)$ 8

$E(MST) = E(MSE) + n\sigma_T^2 \Rightarrow \frac{E(MST) - E(MSE)}{n} = \sigma_T^2 \Rightarrow \frac{E(MST - MSE)}{n} = \sigma_T^2$

$\sigma_T^2 = \frac{MST - MSE}{n}$

هشدار: اگر متغیر باشد، صفر قرار دهید.

تکرار (تغیاری) حافظه است.

$Var(Y_{ij}) = \sigma_T^2 + \sigma^2$

مولفه های واریانس

مولفه های واریانس:

سهم واریانس خطا: $\frac{\sigma^2}{\sigma_T^2 + \sigma^2}$

سهم واریانس تیمارها: $\frac{\sigma_T^2}{\sigma_T^2 + \sigma^2}$

درجه در سهم تیمارها بیشتر نتیجه می گیریم که تیمارها تاثیرگذار روی پاسخ

* MS ها مثبت هستند پس SS ها همیشه مثبت هستند *

« ۱۴۰۳، ۸، ۱۹ »

$$\begin{cases} H_0: \mu_i = \mu_j & \text{اگر قبول نکند (دوطرفه)} \\ H_1: \mu_i \neq \mu_j & \text{اگر قبول کند (دوطرفه)} \end{cases} \quad \begin{cases} H_0: \mu_i > \mu_j & \text{اگر قبول نکند (یکطرفه)} \\ H_1: \mu_i < \mu_j & \text{اگر قبول کند (یکطرفه)} \end{cases}$$

فرض کن که قبول های زوجی دوطرفه هستند. اگر قبول دایره (دوطرفه هست).

تمرین: در طرح تصادفی ساده، مدل با اثرات ثابت شده و مدل با اثرات تصادفی

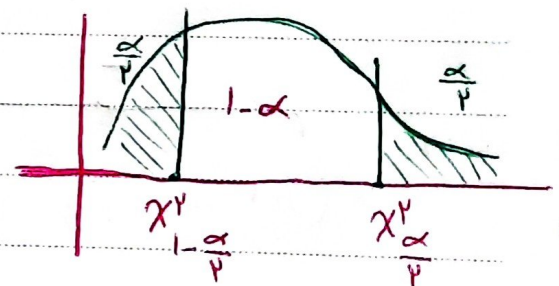
بدست آورده.

طرح تصادفی ساده، مدل با اثرات تصادفی (طرح متقابل)

- فاصله اطمینان برای μ (واریانس خطاها):

$$Q = \frac{(N-a) \text{MSE}}{\mu^2} \sim \chi^2_{(N-a)}$$

کیت مجاری برای μ^2



$$P\left(\chi^2_{1-\frac{\alpha}{p}}(N-a) < Q < \chi^2_{\frac{\alpha}{p}}(N-a)\right) = 1-\alpha \Rightarrow P\left(\frac{(N-a)\text{MSE}}{\chi^2_{\frac{\alpha}{p}}(N-a)} < \mu^2 < \frac{(N-a)\text{MSE}}{\chi^2_{1-\frac{\alpha}{p}}(N-a)}\right) = 1-\alpha \Rightarrow$$

$$\mu^2 = \left(\frac{(N-a)\text{MSE}}{\chi^2_{\frac{\alpha}{p}}(N-a)}, \frac{(N-a)\text{MSE}}{\chi^2_{1-\frac{\alpha}{p}}(N-a)} \right)$$

فاصله اطمینان برای μ^2

- فاصله اطمینان برای مجموع واریانس تیمارها $\left(\frac{\mu^2}{\sigma^2 + \sigma^2_T} \right)$

$$\begin{aligned}
 & \frac{(a-1) MST}{\sigma^2 + n\sigma_T^2} \sim \chi^2_{(a-1)} \quad \dots \quad \frac{(a-1) MST}{\sigma^2 + n\sigma_T^2} \\
 & \frac{(N-a) MSE}{\sigma^2} \sim \chi^2_{(N-a)} \quad \xrightarrow{\text{طبق قضیه تیلور}} \quad \frac{(a-1)}{(N-a)} \sim F(a-1, N-a)
 \end{aligned}$$

$$p\left(F(a-1, N-a) < \frac{MST}{MSE} \cdot \frac{\sigma^2}{\sigma^2 + n\sigma_T^2} < F_{1-\frac{\alpha}{p}}(a-1, N-a)\right) = 1 - \alpha \iff \frac{MSE}{MST}$$

$$p\left(\frac{MSE}{MST} \cdot F_{1-\frac{\alpha}{p}} < \frac{\sigma^2}{\sigma^2 + n\sigma_T^2} < \frac{MSE}{MST} \cdot F_{\frac{\alpha}{p}}\right) = 1 - \alpha \xrightarrow{\text{وارزده باشد}} p\left(\frac{MST}{MSE} \cdot \frac{1}{F_{\frac{\alpha}{p}}} < \frac{\sigma^2 + n\sigma_T^2}{\sigma^2} < \frac{MST}{MSE} \cdot \frac{1}{F_{1-\frac{\alpha}{p}}}\right) = 1 - \alpha$$

$$\xrightarrow{-1 \div n} p\left(\frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{\frac{\alpha}{p}}} - 1\right) < \frac{\sigma_T^2}{\sigma^2} < \frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{1-\frac{\alpha}{p}}} - 1\right)\right) = 1 - \alpha \quad \text{طبق قانون ریاضی}$$

$$p\left(\frac{L}{1} < \frac{\sigma_T^2}{\sigma^2} < \frac{U}{1}\right) \xrightarrow{\text{صورت را با خارج می کنیم}} p\left(\frac{L}{L+1} < \frac{\sigma_T^2}{\sigma^2 + \sigma_T^2} < \frac{U}{U+1}\right) = 1 - \alpha \Rightarrow$$

$$p\left(\frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{\frac{\alpha}{p}}(a-1, N-a)} - 1\right) < \frac{\sigma_T^2}{\sigma^2 + \sigma_T^2} < \frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{\frac{\alpha}{p}}(N-a, a-1)} - 1\right)\right)$$

$$\frac{\sigma_T^2}{\sigma^2 + \sigma_T^2} \sim \left(\frac{L}{L+1}, \frac{U}{U+1}\right)$$

$$F_{1-\frac{\alpha}{p}}(a-1, N-a) = \frac{1}{F_{\frac{\alpha}{p}}(N-a, a-1)} \quad \text{فیش را برعکس کنیم می شود فیش}$$

$F_{\alpha} = (N-a, a-1)$ درجه آزادی را جایگزین کنیم که بجای F_{α} بنویسیم F_{α}

$$P\left(\frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{\alpha}(a-1, N-a)} - 1 \right) < \frac{\frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{1-\alpha}} - 1 \right)}{\frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{1-\alpha}} - 1 \right)} = 1 - \alpha\right)$$

مثال ۳-۸ کتاب صفحه ۹۸: نکته سوال: تیمارها به تصادف انتخاب شده اند. که در صورت سوال گفته تیمارها

به تصادف انتخاب شده اند.

عامل: کشتی ماهیگیری ها هدف: بررسی کشتی (یعنی یکسانی در تولید) ماهی ها بافت در این کارخانه از نظر اصناف کشتی

بالش

هدف بررسی تیمار است

تیمار (انواع) اصطلاح ماهی ها } ماهی ۱
ماهی ۲
ماهی ۳
ماهی ۴

مدل با اثرات تصادفی شده. چرا؟ چون تیمارها به تصادف انتخاب شده اند (چون صورت سوال گفته ماهی ها به تصادف

انتخاب شده است)

تکرارها یکسان است پس می توانیم طرح متقابل

تعریف است

مدل طرح ۸: $y_{ij} = \mu + T_i + \epsilon_{ij}$

شرایط مدل با اثرات تصادفی ۹: $\epsilon_{ij} \sim N(0, \sigma^2)$

ناحیه بالشت $\leftarrow \epsilon_{ij}$ و T_i

چون T_i انتخابی است، $T_i = 0$ نیست و T_i باید توزیع داشته باشد $T_i \sim N(0, \sigma^2)$

یعنی با احتمال ۱ $T_i = 0$ همیشه

فرضیه: $H_0: \sigma^2 = 0$
 $H_1: \sigma^2 > 0$

آزمون فاریس

$$SST_0 = \sum_i \sum_j y_{ij}^2 - \frac{Y_{..}^2}{N} = 119,95$$

$$SST = \sum_{i=1}^a \frac{y_i^2}{n} - \frac{Y_{..}^2}{N} = 89,19$$

$$SSE = SST_0 - SST = 22,75$$

جدول ANOVA

Source	df	SS	MS	F ₀
treatment (ماشین ها)	3	89,19	$MST = \frac{89,19}{3} = 29,73$	$F_0 = \frac{MST}{MSE} = 15,48$
Error	$14 - 3 = 11$	22,75	$MSE = \frac{22,75}{11} = 2,07$	
Total	14	119,95		

$$F_{0.05}(3, 11) = 3,1 \Rightarrow F_0 > F_{\alpha} \Rightarrow RH$$

نتیجه آزمون آماری: با اطمینان ۹۵٪ بر اساس نتایج موجود تفاوت معنی داری بین تیمارها وجود دارد.

نتیجه آزمون کلی: با اطمینان ۹۵٪ بر اساس نتایج موجود تفاوت معنی داری بین کارخانه ها وجود دارد.

نکته مهم: یعنی تعیین داریم به کل ماشین های کارخانه

نکته: در حالت مدل با اثرات تصادفی ما جوابیات زنی نداریم. (معنی ندارد)

مولدهای داربالتس را برآورد کنید و سهم هر یک را محاسبه و تحلیل کنید و فاصله اطمینان بیابید.

۱۳۵۳/۸/۲۰

Subject :

Date

تحلیل و آمار در جوامع و کسب

طرح تعادلی ساده - طرح متقابل - مدل با اثرات تعادلی

مؤلفه‌های واریانس

$$\text{var}(y_{ij}) = \sigma^2 + \sigma^2_T$$

$$\begin{cases} E(MST) = \sigma^2 + n\sigma^2_T \\ E(MSE) = \sigma^2 \end{cases} \quad \text{طرح متقابل}$$

$$\hat{\sigma}^2_T = MSE = 1,9 \quad \text{از جدول ANOVA}$$

$$\hat{\sigma}^2_T = \frac{MST - MSE}{n} = \frac{291,73 - 1,9}{4} = 4,94$$

برآورد نقطه‌ای مؤلفه‌های واریانس

برآورد سهم مؤلفه‌های واریانس

$$\text{سهم واریانس تیمارها} = \frac{\hat{\sigma}^2_T}{\hat{\sigma}^2 + \hat{\sigma}^2_T} = \frac{4,94}{1,9 + 4,94} = \frac{4,94}{6,84} = 0,72$$

برآورد سهم مؤلفه‌های واریانس

$$\text{آزاد سهم واریانس خطاها} = \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \hat{\sigma}^2_T} = 1 - 0,72 = 0,28$$

(تفسیر پذیری خطاها)

تحلیل: ۷۸ درصد از تغییرات پاسخ توسط تیمارها بیان می‌شود و ۲۲ درصد باقی مانده توسط خطاها بیان می‌شود.

نتیجه: بیشتر تغییر پذیری کل (پاسخ) توسط تیمارها بیان می‌شود. یعنی تیمارها روی پاسخ تأثیرگذارند.

$$\hat{\sigma}^2 = \left(\frac{(N-a)MSE}{\chi^2_{\frac{\alpha}{p}}(N-a)} \right) \text{ و } \left(\frac{(N-a)MSE}{\chi^2_{1-\frac{\alpha}{p}}(N-a)} \right) = \left(\frac{12 \times 1,9}{\chi^2_{0,05}(14)} \text{ و } \frac{12 \times 1,9}{\chi^2_{0,95}(14)} \right) = (0,98 \text{ و } 5,18)$$

۲۳,۳۴ ۶,۲

تحلیل: با اطمینان ۹۵ درصد واریانس خطاها بین ۰,۹۸ تا ۵,۱۸ تغییر می‌کند.

* حاصله اطمینان اطلاعات بیشتری به ما می‌دهد *

فاصله اطمینان برای سهم داربالتی (تیارها):

$$\alpha = 0.05 \quad F_{0.05}(3, 12) = 4.147$$

$$F_{0.05}(11, 3) = 5.122$$

$$\frac{\omega_T^2}{\omega_T^2 + \omega_T^2} = \left(\frac{L}{L+1}, \frac{U}{U+1} \right) = \left(\frac{0.425}{0.425+1}, \frac{20.21}{20.21+1} \right) = (0.38, 0.95)$$

$$L = \frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{1}{F_{\frac{\alpha}{p}}(a-1, N-a)} - 1 \right) = \frac{1}{8} \left(\frac{29.73}{1.9} \cdot \frac{1}{4.147} - 1 \right) = 0.425$$

$$U = \frac{1}{n} \left(\frac{MST}{MSE} \cdot \frac{F_{\frac{\alpha}{p}}(N-a, a-1)}{p} - 1 \right) = \frac{1}{8} (15.28 \times 5.122 - 1) = 20.21$$

تحلیل: با اطمینان ۹۵٪ سهم تصیری (تیارها) بین ۱۹۵ تا ۲۳۸ از کل می باشد.

فصل ۴: بررسی تفاوت مدل

بررسی شرایط مدل را با نرم افزار مینی تاب (Mini Tab) انجام می دهیم.

ابتدا باید باقی مانده ها را به دست آوریم.

$$y_{ij} = \mu + T_i + \epsilon_{ij}$$

$$\hat{\epsilon}_{ij} = e_{ij} = y_{ij} - \hat{y}_{ij}$$

- ۱- تعریف مدل خطاها
- ۲- هم واریانس خطاها
- ۳- ناهمبستگی خطاها

بررسی وجود باقی مانده مشترک: $|e_{ij}^*| = \frac{|e_{ij}|}{\sqrt{MSE}} > 3$ (لیست داده ها برت) هستند.

نرم نمودار عیبی برای باقی مانده ها ← وجود باقی مانده های برت.

آزمون بارتلت (در آزمون میانیتم و باید تمام مباحث آماری به روش و تمرینات مباد

Subject :
Date

$$\left\{ \begin{array}{l} H_0: p_1 = \dots = p_a \\ H_1: \text{Not } H_0 \end{array} \right.$$

آزمون بارتلت برای بررسی هم وابستگی خطاها

$$X^2_0 = 21.3029 \times \frac{q}{c} \sim X^2_{(a-1)} \quad q = (N-a) \log_{10} S_p^2 - \sum_{i=1}^a (n_i-1) \log_{10} S_i^2$$

$$C = 1 + \frac{1}{3(a-1)} \left(\sum_{i=1}^a (n_i-1) - (N-a) \right)$$

$$S_p^2 = \frac{\sum_{i=1}^a (n_i-1) S_i^2}{N-a}$$

نکته: آزمون بارتلت زمانی که می شود که خطاها تکرار باشند.

آزمون بارتلت برای مثال ۱-۳: $X^2_0 = 0.693$ و $X^2_{(2)} = 9.210$ نتیجه می گیریم چون $X^2_0 < X^2_{(2)}$ فرض H_0

پذیرفته می شود یعنی وابستگی ها برابرند.

برای ناهمبستگی آزمون خلی و جوردان

آزمون دو رتبه - واسون: خود همبستگی را بررسی می کند.

$$\left\{ \begin{array}{l} H_0: p_1 = 0 \\ H_1: p_1 \neq 0 \end{array} \right.$$

آزمون نزدیک ۲ باشد، خود همبستگی ۰ است. یعنی خطاها ناهمبسته هستند.

این آزمون با DW یا p-value می دهد.

معمولا تبدیلات نرمال (سازو هم دارایی) تقریبا همیشه است $\Leftarrow$ مثل تبدیل دالگس - کالکس

تبدیل جبر $\Leftarrow$ تبدیل جذر + 1 $\Leftarrow$ تبدیل $\sqrt{\text{Arc sin}}$ ، تبدیل لگاریتمی با سطح.

تبدیل رانسی $\Leftarrow$ رانسی هانی زنیسم.

در تبدیلات box cox $\Leftarrow$

* y_{ij} $\left\{ \begin{array}{l} \text{فرم ساده} \\ y_{ij}^{\lambda} \quad \lambda \neq 0 \\ \log(y_{ij}) \quad \lambda = 0 \end{array} \right.$

$\left\{ \begin{array}{l} \text{فرم پیچیده} \\ \frac{y_{ij}^{\lambda} - 1}{\lambda \bar{y}} \quad \lambda \neq 0 \\ \log(y_{ij}) \times \bar{y} \quad \lambda = 0 \end{array} \right.$

$\bar{y}$ میانگین هندسی با سطح ها

« ۱۴۳۱/۸/۲۶ »

- ۱- خطاها نرمال باشند $\Leftarrow$ اگر نرمال نباشند
- ۱- بافتن تبدیل نرمال ساز (تبدیلات دالگس - کالکس)
- ۲- روش های نظایر بقوی (کرنجول) که در شکل دالگس ص ۸۸ کتاب برنولدی

شرایط مدل

۱- استناد (در تبدیلات) معروف

۲- خطاها هم دارایی باشند $\Leftarrow$ اگر هم دارایی نباشند $\Leftarrow$ بافتن تبدیلات بلانداریناری دارایی

۲- استناد از تعادل فضای ریداست که پس واریانس تیمارها و میانگین آن ها وجود دارد

آزادی آکوستیک دالگس و رانسی ای که هم در شرایط هم دارایی و

ناهمبستگی برقرار باشند و فقط شرط نرمال بودن برقرار نیست.

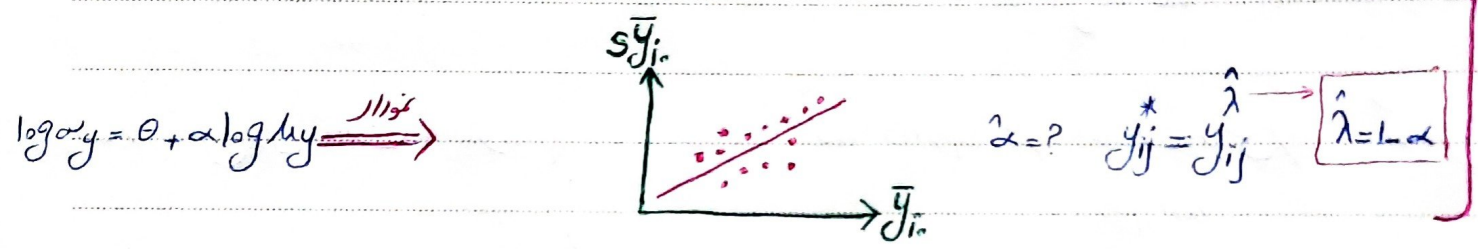
- ① $y^* = \log(y)$
- ② $y^* = \text{Arcsin} \sqrt{y}$
- ③ $y^* = \sqrt{y} - \sqrt{y+1}$
- ④ $y^* = \frac{1}{y}$
- ⑤ تبدیلیات کالکس - کالکس

۱- استاندارد از تبدیلیات معروف

نشیب خط
 $\log \sigma_y = \theta + \alpha \log \mu_y$

۲- استاندارد از مقدار نه فعلی زیر

از این استاندارد
 $\log \sigma_y \propto \log \theta + \alpha \log \mu_y$
 متناسب
 $\propto \theta \mu_y$
 ممکن است یک رابطه به این صورت باشد



اگر $\alpha = 1$ شد، تبدیل لگاریتمی می گیریم
 $\lambda = 0 \leftarrow y_j^* = \log y_j$

نمونه مثال ۴۲ ص ۱۲۲ کتاب رادیشی و ماژیم اثر را می دهیم.

درصد بنه	۱	۲	۳	۴	۵
۱۵٪	۷	۷	۱۵	۱۱	۹
۲۰٪	۱۲	۱۷	۱۲	۱۸	۱۸
۲۵٪	:	:	:	:	:
۳۰٪	:	:	:	:	:
۳۵٪	:	:	:	:	:

« قسمت کار با بنه اثر از منحنی تب »

مثال ۳-۱ درصد بنه ۵

Factor treatment	باز
دوستون در منحنی تب	ماژیم کنیم

متغیرهای ما در اینجا درصد بنه های و متناوب کشش (با بنه) هستند.

در قسمت graph $\leftarrow$ box plot (رسم کریم) $\leftarrow$ این که جعبه اش (کوچکتر و دریاش (بزرگ) کمتر.
آنگاه جدا در یک سطح باشند یعنی تفاوت ندارند اما اینها در یک رانجا باشند.

در stat در ANOVA همکاری داریم اینجا ای کی (هم).
stat $\rightarrow$ ANOVA $\rightarrow$ General linear model

Generalized linear model $\Rightarrow$ (در SPSS جدا کرده)
تحقیق و پویش جعبه ها هستند اما در قسمت (رسم کریم) در SPSS تب هستند.

عمود اثرات اصلی $\leftarrow$ (همه میانی را نشان می دهد یعنی چه تغییری در میانی به وجود آمده)
stat $\rightarrow$ main Effect plot $\leftarrow$

stat $\rightarrow$ ANOVA $\rightarrow$ one way

مقایسات زوجی فصل کوکی - رانیت - فیشتر $\leftarrow$
comparisons $\Rightarrow$ مقایسات زوجی $\Rightarrow$ (در آزمون) انجام می دهد.

در قسمت graph $\leftarrow$ for in one رانیک بریم و هر یک از اول را هم بریم.

storage (یعنی به چیزی) در داده ها همان ذخیره کند $\leftarrow$

Graph $\Rightarrow$ probability plot $\Rightarrow$ (آزمون نرمال بودن یا آرایش عددی)

اینجا residual ما را وارد می کنیم و نمودار بهمان می دهد.

stat $\rightarrow$ ANOVA $\rightarrow$ test for equal variance $\Rightarrow$ response $\Rightarrow$ residual (داریم)
Factor $\Rightarrow$ treatment (داریم)

که شکل نشان می دهد تقریباً دریاش (ها یکی است).

بافتن کاغذیں خطائی شود ← بلوک بندی - تصانیف کردن
 باعث شناسایی خطائی شود ← تکرار

Subject: _____
 Date: _____

۱۳۰۳، ۹، ۳

فصل پنجم: طرح بلوکی

طرح بلوکی (کامل): ۱- یک طرح تک عاملی است یعنی هدف بررسی سطوح یک عامل است.

۲- یک عامل محدود کننده (بلوک) در تصانیف کردن وجود دارد. ۳- بقیه شرایط یکسان و فقط تیمارها متفاوت اند.

۴- بلوکی کامل بالشت (یعنی در هر بلوک همه تیمارها به کار رفته باشند در غیر این صورت بلوکی ناقص می شود)

تیمار (سطوح عامل اصلی)
 تیمار ۱، تیمار ۲، ...

بلوک ۱	A	C	B	D	E
بلوک ۲			B	B	
بلوک ۳					
بلوک ۴					

خطاهای آزمایش ← اثر بلوک ز ام ← اثر تیمار ز ام

$$y_{zj} = \mu + T_i + \beta_j + \epsilon_{zj}$$

مدل طرح ۳: ϵ_{zj} میانگین کل

پانچ تیمار ز ام در بلوک ز ام

بلوک ها (Block)

	۱	۲	...	b
۱	y_{11}	y_{12}	...	y_{1b}
۲	y_{21}	y_{22}	...	y_{2b}
...	...	...	...	...
a	y_{a1}	y_{a2}	...	y_{ab}

$i = 1, 2, \dots, a$
 $j = 1, 2, \dots, b$
 $N = ab$

مدل با اثرات تثبیت شده است در طرح بلوکی

شرایط مدل ۳: $\epsilon_{zj} \sim iid N(0, \sigma^2)$

$\sum_{i=1}^a T_i = 0$ و $\sum_{j=1}^b \beta_j = 0$

میانگین کل
 $T_i = \bar{y}_{i.} - \bar{y}_{..}$

اثر تیمار

میانگین کل
 $\beta_j = \bar{y}_{.j} - \bar{y}_{..}$
 میانگین بلوکی

فرضیه: $H_0: T_i = 0$ و $H_1: \text{Not } H_0$

برای آزمون کمر (فرضیه) آمار واریانس انجام می‌دهیم.
(آمار واریانس (ANOVA) یعنی تجزیه واریانس کل به مولفه‌های
موجود در مدل)

$$SSTo = SST + SSB + SSE$$

$$df \Rightarrow N-1 = a-1 + b-1 + (a-1)(b-1)$$

فرضیه محاسباتی SS ها:

$$SSTo = \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$SST = \sum_{i=1}^a \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{N}$$

$$SSB = \sum_{j=1}^b \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{N}$$

$$SSE = \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2 - \sum_{i=1}^a \frac{y_{i.}^2}{b} - \sum_{j=1}^b \frac{y_{.j}^2}{a} + \frac{y_{..}^2}{N} = SSTo - SST - SSB$$

امید $E(MS)$ ها: برای تعیین F ها نیاز است (در مدل با اثرات تبیین شده)

$$E(MST) = \sigma^2 + \frac{b \sum_{i=1}^a T_i^2}{a-1}$$

$$E(MSB) = \sigma^2 + \frac{a \sum_{j=1}^b \beta_j^2}{b-1}$$

$$E(MSE) = \sigma^2$$

آزمون برای بلوک ای می دهیم. اما در تمام افراد هر دو را به هم می دهد.
اثر متقابل بین بلوک و تیمار وجود ندارد.

Subject :
Date :

تحت فرض H_0 کدام دو $E(MS)$ ها یکی می شوند؟
 $E(MST) \stackrel{H_0}{=} E(MSE)$

$$F_0 = \frac{MST}{MSE} \stackrel{H_0}{=} F(a-1, (a-1)(b-1))$$

$$F_0 > F_{\alpha}(a-1, (a-1)(b-1)) \Rightarrow R H_0$$

مثال ۱-۵ صفحه ۱۲۲ کتاب: آزمایش آزمونی سختی یک فلز بر اساس ۴ طبقه را بررسی کنیم. از هر طبقه ۳ تیر داریم. در اینجا خود مسئله گفته که یک طرح بلوکی اجرا شده است.

(تفسیر پذیری بین قطعات متفاوت به نی طرح بلوکی) قطعه آزمایش					
تیمار	۱	۲	۳	۴	$\sum y_{i.}$
۱	۹,۳	۹,۴	۹,۶	۱۰	۳۸,۳
۲	۹,۴	۹,۳	۹,۸	۹,۹	۳۸,۴
۳	۹,۲	۹,۴	۹,۵	۹,۷	۳۷,۸
۴	۹,۷	۹,۶	۱۰	۱۰,۲	۳۹,۵
$\sum y_{.j}$	۳۷,۴	۳۷,۷	۳۸,۹	۳۹,۸	$154 = \sum$

کدگذاری: یعنی صفای یک عددی کنیم که شماره شود (مثلاً صفای ۱۰ x ۱۰ کنیم) یا ۱۰ x ۹ (همه صفای ۱۰ x ۹ کنیم)

توضیح داده شود چرا بلوکی است؟ ۱- عامل اصلی ۲- عامل محدود کننده بلوک

$$\text{مدل طرح: } y_{ij} = \mu + T_i + \beta_j + \varepsilon_{ij}$$

شرایط مدل: اثرات ثابت شده $\sum \beta_j = 0$ و $\sum T_i = 0$ و $\varepsilon_{ij} \sim N(0, \sigma^2)$ iid

$$\begin{cases} H_0: T_i = 0 \quad \forall i \\ H_1: N_0 + H_0 \end{cases}$$

$$SST_0 = \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2 - \frac{y_{..}^2}{N} = 1,29$$

$$SST = \sum_{i=1}^a \frac{y_{i.}^2}{b} - \frac{y_{..}^2}{N} = 0,385$$

$$SSB = \sum_{j=1}^b \frac{y_{.j}^2}{a} - \frac{y_{..}^2}{N} = 0,125$$

$$SSE = 1,29 - 0,125 - 0,385 = 0,78$$

جدول ANOVA

منابع تغییرات Source	df	SS	MS	F ₀
treatment تیمار	3	0,385	$MST = \frac{0,385}{3} = 0,128$	
Block بلوک	3	0,125		$F_0 = \frac{0,128}{0,089} = 1,438$
Error خطا	9	0,78	$MSE = \frac{0,78}{9} = 0,086$	$F_{0,05}(3,9) = 3,18$
Total کل	15	1,29		

$$F_0 > F_{\alpha} \Rightarrow R.H.$$

نتیجه گیری: آماره F با اطمینان 95 درصد بر اساس شواهد موجود اختلاف معنی داری بین تیمارها وجود دارد.

نتیجه گیری: کل با اطمینان 95 درصد تفاوت معنی داری بین این انواع نیتروژن برای اسانس صیقلی.

Subject:

Date:

ANCOVA طرح تصادفی (تصادفی)

Source	df	SS	MS	F ₀
treatment	3	0.385	0.13	$F_0 = \frac{0.13}{0.075} = 1.73$
Error	12	0.905	$\frac{0.905}{12} = 0.075$	
Total	15	1.29		$F(3, 12) = 3.19$

نوع تیغه ها تفاوت معنی داری از نظر آماری یعنی تفاوت دارند $\Rightarrow A H_0 \Rightarrow F_0 \neq F_{\alpha}$

آزمون تی (تست) که خطای آن زیاد باشد، ممکن است روی نتیجه ها اثر بلند از قبل اینجا

مقایسات زوجی انجام دهید که نوع تیغه ها با هم متفاوت اند گروه بندی کنید. کدام تیغه معتبر است؟

$$\bar{Y}_1 = 9.575 \quad \bar{Y}_2 = 9.4 \quad \bar{Y}_3 = 9.45 \quad \bar{Y}_4 = 9.175 \quad \bar{Y}_5 = 9.175 \quad \bar{Y}_6 = 9.175$$

« 143 / 9 / 10 »

مقایسات زوجی، روش دانکن چون طرح بلوکی از MSE طرح بلوکی استفاده کنیم.

$$R_m = t_{\alpha}(m, f) \sqrt{\frac{MSE}{b}}$$

$$MSE = 0.00889$$

$$m = (\text{اختلاف رتبه } \bar{Y}_i \text{ و } \bar{Y}_j) + 1$$

$$b = 2$$

	3	1	2	4
3		R_1	R_2	R_4
1	0.125		R_1	R_2
2	0.15	0.025		R_1
4	0.425*	0.3*	0.275*	

3				
1	AH ₀			
2	AH ₀	AH ₀		
4	RH ₀	RH ₀	RH ₀	

فرض کنید $E(y_{ij}) = \mu + T_i + \beta_j$ از آن رو می‌توانیم وضرب را تبدیل به جمع می‌کنیم (مدل را از حالت

عکس می‌تبدیل به خطی می‌کنیم)

$$LSD = R_p = \sqrt{\frac{0.000189}{4}} = 0.15085 \quad R_p = 0.157 \quad R_f = 0.1412$$

گروه اول	گروه دوم
$\bar{y}_{1.}$	$\bar{y}_{2.}$
$\bar{y}_{1.}$	
$\bar{y}_{1.}$	

داده‌های گمشده طرح بلکی کامل: missing Data

	block(j)						
	1	2	...	b		$y_{i.}$	
Treatment(i)	1	y_{11}	y_{12}	...	y_{1j}	y_{1b}	$y_{1.}$
	2	y_{21}	y_{22}	...	y_{2j}	y_{2b}	$y_{2.}$
	...	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	i	y_{i1}	y_{i2}	...	y_{ij}	y_{ib}	$x + y'_{ij}$
	a	y_{a1}	y_{a2}	...	y_{aj}	y_{ab}	
$y_{.j}$	$y_{.1}$	$y_{.2}$	...	$(x + y'_{ij})$	$y_{.b}$	$y_{..}$	

هر تعداد داده گمشده که داریم باید به تعداد آن از داده‌های آزادی از خطا و داده‌های آزادی کم می‌شود.

Min SSE(x): روش اصلی

$$SSE = \sum_i \sum_j y_{ij}^2 - \sum_i \frac{y_{i.}^2}{b} - \sum_j \frac{y_{.j}^2}{a} + \frac{y_{..}^2}{ab} = y_{11}^2 + \dots + x^2 + \dots + y_{ab}^2 - \frac{y_{1.}^2 + \dots + (x + y'_{ij})^2}{b}$$

برای پایانتار کردن و طرح داد و وقت نهالینم برای یک طرح مقایسه زوجی اینجا داریم که بدانید بایدیم

Subject:

Date

$$\frac{y_{.j}^2 + \dots + (x + y'_{.j})^2 + \dots}{a} + \frac{(x + y'_{..})^2}{ab}$$

نسبت می گیریم نسبت به x SSE مشتق می گیریم نسبت به x $\Rightarrow \frac{\partial SSE(x)}{\partial x} = 2x - 2 \frac{(x + y'_{.j})}{b} - 2 \frac{(x + y'_{.j})}{a} + 2 \frac{(x + y'_{..})}{ab} = 0$

مشتق و $\frac{\partial}{\partial x}$ $\Rightarrow abx - ax - ay'_{.j} - bx - by'_{.j} + x + y'_{..} = 0 \Rightarrow x(ab - a - b + 1) = ay'_{.j} + by'_{.j} - y'_{..}$

نسبت می گیریم نسبت به x $\Rightarrow x = \frac{ay'_{.j} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = y'_{.j}$

داره امشده جایگذاری می شود پس در داده جدول ANOVA طرح بلوک کامل قبل محاسبه می شود تنها

وقت کند یک درجه آزادی از خطا سری شود (یعنی نه آزادی هر مقدار داده امشده که برگردانده از درجه آزادی خطاها سری کرد)

و به همان نسبت از درجه آزادی کل سری شود.

تمرین: فرض کنید دو داده امشده داریم برای همین مثال حل شود.

۱۳۰۳/۹/۱۸

هر جا F جدول ANOVA دارد شده باید مقایسات زوجی اینجا داریم (فقط در طرح های نسبت شده اینجا)

می شود. در طرح تصادفی ساده اینجا نمی شود

طرح تصادفی ساده نسبت شده نامعادل

$$y_{ij} = \mu + T_i + \epsilon_{ij}$$

$$\epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2) \text{ و } \sum_{i=1}^a n_i T_i = 0$$

شرایط مدل

$$E(MST) = E\left(\frac{SSE}{N-a}\right) = \frac{1}{N-a} E(SSE) = \frac{1}{N-a} E\left(\sum_{i=1}^a \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^a \frac{y_{i.}^2}{n_i}\right) = \frac{1}{N-a} \left[\sum_{i=1}^a \sum_{j=1}^{n_i} E(y_{ij}^2) - \frac{E(y_{i.}^2)}{n} \right]$$

$$E(y_{ij}^2) = E(\mu + T_i + \epsilon_{ij})^2 = E(\mu^2 + T_i^2 + \epsilon_{ij}^2 + 2\mu T_i + 2\mu \epsilon_{ij} + 2T_i \epsilon_{ij}) = \mu^2 + T_i^2 + \sigma^2 + 2\mu T_i + 0 + 0$$

$$= \sum_{i=1}^a \sum_{j=1}^{n_i} E(y_{ij}^2) = N\mu^2 + \sum_{i=1}^a n_i T_i^2 + N\sigma^2 + 2\mu \sum_{i=1}^a n_i T_i$$

$$E(y_{i.}^2) = E\left(\sum_{j=1}^{n_i} y_{ij}\right)^2 = E\left[\left(\sum_{j=1}^{n_i} \mu + T_i + \epsilon_{ij}\right)^2\right] = E(n_i \mu^2 + n_i T_i^2 + \epsilon_{i.}^2 + 2n_i \mu T_i + 2n_i \mu \epsilon_{i.} + 2n_i T_i \epsilon_{i.}) = n_i \mu^2 + n_i T_i^2 + n_i \sigma^2 + 0 + 0 + 0$$

$$\frac{E(y_{i.}^2)}{n_i} = n_i \mu^2 + n_i T_i^2 + \sigma^2 + 2n_i \mu T_i = \sum_{i=1}^a \frac{E(y_{i.}^2)}{n_i} = N\mu^2 + \sum_{i=1}^a n_i T_i^2 + a\sigma^2 + 0 = \left[N\mu^2 + \sum_{i=1}^a n_i T_i^2 + N\sigma^2 - N\mu^2 - \sum_{i=1}^a n_i T_i^2 - a\sigma^2 \right] \frac{1}{N-a} = \frac{1}{N-a} [a\sigma^2(N-a)] = \sigma^2$$

$$N\sigma^2 - N\mu^2 - \sum_{i=1}^a n_i T_i^2 - a\sigma^2 \left[\frac{1}{N-a} \right] = \frac{1}{N-a} [a\sigma^2(N-a)] = \sigma^2$$

حل بالترتيب تعاقبي سادس

$$y_{ij} = \mu + T_i + \epsilon_{ij}$$

$\epsilon_{ij} \sim \text{iid } N(0, \sigma^2)$
 $T_i \sim \text{iid } N(0, \sigma_T^2)$

$$E(MST) = \frac{1}{a-1} E\left[\sum_{i=1}^a \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N}\right]$$

$$E(y_{i.}^2) = n\mu^2 + n\sigma_T^2 + n\sigma^2$$

$$\frac{E(y_{i.}^2)}{n} = a\mu^2 + \sigma^2$$

$$E(y_{i.}^2) = E\left(\sum_{j=1}^n y_{ij}\right)^2 = E\left[\sum_{j=1}^n (\mu + T_i + \epsilon_{ij})\right]^2 = E(n\mu^2 + nT_i^2 + \epsilon_{i.}^2 + 2n\mu T_i + 2n\mu \epsilon_{i.} + 2nT_i \epsilon_{i.}) = n\mu^2 + nT_i^2 + n\sigma^2 + 0 + 0 + 0$$

$$PAPCO \quad n\mu^2 + nT_i^2 + n\sigma^2 = n\mu^2 + n\sigma_T^2 + n\sigma^2$$

$$\sum_{i=1}^a \frac{E(y_{i.}^2)}{n} = a\mu^2 + a\sigma_T^2 + a\sigma^2$$

$$E(T_i \epsilon_{i.}) = E(T_i) E(\epsilon_{i.})$$

$$\frac{1}{a-1} \left[a n \mu^2 + a n \sigma_T^2 + a \sigma^2 - a n \mu^2 - a \sigma^2 - a n \mu^2 - n \sigma_T^2 - \sigma^2 \right] = \frac{\sigma_T^2 n (a-1) + \sigma^2 (a-1)}{a-1} = n \sigma_T^2 + \sigma^2$$

$$\sum \sum y_{ij} = N \mu + n \sum T_i + \epsilon_{..} = N \mu^2 + n \sum T_i^2 + \epsilon_{..}^2 + \dots = n \mu^2 + n a \sigma_T^2 + N \sigma^2 = N \mu^2 + n a \sigma_T^2 + N \sigma^2$$

$$\sum_{i=1}^a \sum_{j=1}^{n_i} k = k \sum_{i=1}^a n_i = N k$$

$$n_1 + n_2 + \dots + n_a = N$$

$$\sum_{i=1}^a n_i = N$$

۱۴۰۳ / ۹ / ۲۴

طرح مربع لاتین ۸

۱- یک طرح تک عاملی است. ۲- دو عامل محدود کننده در تعدادی که در آن وجود دارد. ۳- عقبه شرایطی بسیار

و فقط تیارها متفاوت اند. ۴- تیارها به تعارف روی واحدهای آزمایشی اعمال می شوند. ۵- طرح بالاس

(یعنی تعداد سطوح عوامل محدود کننده و تعداد تیارها برابرند) تعداد تیارها p

مثال: فرض کنید می خواهیم چهار نوع لایتنر را از نظر میزان ساییدگی بررسی کنیم. دو عامل محدود کننده فرض کنید

وجود دارد و نوشته ایم شرایط برای این دو عامل بسیار کنیم.

چهار نوع خود دارد و چهار موقعیت طرح ها $\leftarrow$

- ۱- جلوراست
- ۲- جلوچپ
- ۳- عقب راست
- ۴- عقب چپ

عامل محدود کننده اول عامل محدود کننده دوم

حروف لاتین (شان) دهنده (نوع تیمارها است).

مثال: طرح مربع 3×3 نوع استاندارد $\leftarrow$

صفون (عوامل مورد اندازه گیری)

	۱	۲	۳
۱	A	B	C
۲	B	C	A
۳	C	A	B

نوع استاندارد $\rightarrow$ یعنی به ترتیب حروف الفبا هم شماره دهیم

مستقل است.

نوع غیر استاندارد $\leftarrow$

	۱	۲	۳
۱	B	A	C
۲	A	C	B
۳	C	B	A

نوع خوار و (k)

طرح مربع لاتین 4×4

$$N = p \times p$$

	۱	۲	۳	۴
۱	A = y_{11k}	B	D	C
۲	B _i	C _i	A	D = y_{24k}
۳	C	D	B	A
۴	D	A	C	B

نوع لاتین $\rightarrow$

تیمار $\leftarrow P = 1, 2, \dots, p$

Row $\leftarrow P = 1, 2, \dots, p$

Col $\leftarrow P = 1, 2, \dots, p$

* طرح مربع 2×2 داریم چون (دو تیمار آزادی) (خطا ۱ می شود) (دو تیمار آزادی) (کاهش ندارد)

اثر عامل محدود کننده سطری

خطاهای آنالیز

مدل طرح :

$$y_{ijk} = \mu + T_i + \beta_j + \gamma_k + \varepsilon_{ijk}$$

اثر عامل ستونی

اثر تکرار

با انتخاب تکرار ۱۱ در ۳۱ سطح عامل سطری در ۳۱ سطح عامل ستونی

شرایط مدل : $\varepsilon_{ijk} \sim N(0, \sigma^2)$ iid

$$\sum_{i=1}^p T_i = 0, \sum_{j=1}^p \beta_j = 0, \sum_{k=1}^p \gamma_k = 0$$

در مدل با اثرات تنبیه شده :

برای آزمون فرضیه فوق : آنالیز واریانس (ANOVA) :

$$SSTo = SST + SS_{Row} + SS_{col} + SSE$$

$$df : N = p \times p = p^2 = (p-1) + (p-1) + (p-1) + (p-2)(p-1)$$

تجزیه : نحوه تجزیه واریانس

$$SSE \Rightarrow N - 1 - 3(p-1) = p^2 - 1 - 3(p-1) = (p-1)(p+1) - 3(p-1) = (p-1)(p+1-3) = (p-1)(p-2)$$

$$\begin{cases} H_0 : T_i = 0 & i=1, \dots, p \\ H_1 : \text{Not } H_0 \end{cases}$$

فرضیات مدل :

$$N = p^2$$

فرمول های محاسباتی SS ها :

$$SSTo = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$$

Subject:

Date

جمع تیارها را باید جدول جدا بنویسیم و در دست تیار رسم

$$SST = \sum_{i=1}^P \frac{Y_{i..}^2}{P} - \frac{Y_{...}^2}{N}$$

$$SS_{Row} = \sum_j \frac{Y_{.j.}^2}{P} - \frac{Y_{...}^2}{N}$$

$$SS_{Col} = \sum_k \frac{Y_{...k}^2}{P} - \frac{Y_{...}^2}{N}$$

تیار	$Y_{i..}$
A	$Y_{1..}$
B	$Y_{2..}$
C	$Y_{3..}$
D	$Y_{4..}$

همه A ها جمع شوند

همه B ها جمع شوند

همه C ها جمع شوند

همه D ها جمع شوند

$$SSE = SST - SS_{Row} - SS_{Col}$$

انید $E(MS)$ ها :

$$E(MST) = \sigma^2 + \frac{P \sum_{i=1}^P T_i^2}{P-1}$$

$$E(MS_{Row}) = \sigma^2 + \frac{P \sum_j B_j^2}{P-1}$$

$$E(MS_{Col}) = \sigma^2 + \frac{P \sum_k \gamma_k^2}{P-1}$$

$$E(MSE) = \sigma^2$$

تحت فرض $H_0: T_i = 0$ که دو امید ریاضی یکی می شوند؟

$$E(MST) \stackrel{H_0}{=} E(MSE) \Rightarrow F_0 = \frac{MST}{MSE} \stackrel{H_0}{\sim} F(p-1, (p-2)(p-1))$$

$$RH: F_0 > F_{\alpha}(p-1, (p-2)(p-1))$$

جدول ANOVA :

source	df	SS	MS	F _o
treatment	p-1	SST	$MST = \frac{SST}{p-1}$	$F_o = \frac{MST}{MSE}$
Row	p-1	SSRow		
col	p-1	SSCol		
Error	(p-2)(p-1)	SSE	$MSE = \frac{SSE}{(p-2)(p-1)}$	$RH_o : F_o > F_{\alpha}(p-1, (p-2)(p-1))$
Total	N-1	SSTo		

نکته : اگر در جدول ANOVA فرض H₀ را نپذیریم (لازم است مقایسات زوجی انجام شود).

تمرین : در طرح مربع لاتین یک داده گسسته را انجام دهید.

تمرین : در طرح مربع لاتین برای دو داده گسسته انجام دهید.

تمرین : یک داده گسسته در طرح مربع لاتین :

جمع تیمار را از جدول داده گسسته

$$y_{ijk} = \frac{p(y'_{i..} + y'_{.j.} + y'_{..k}) - 2y'_{...}}{(p-2)(p-1)}$$

تمرین : فرض داده گسسته اثبات شود.

تمرین : سه مورد با بارهای مختلف به روش حداقل مربعات در دست آورید.

تمرین : یک تمرین از طرح مربع لاتین حل شود مثال ۵-۵ داده ها را ۱۰ + کنیم سپس ۲ × کنیم و این را رسم کنیم.

از داده کشیده باید یکی از درجه آزادی Error و یکی از درجه آزادی total کم شود.

مثال ۲: ۴۵: مسئله ۵ نوع فرمول بندی (دنیامیت) را می توانیم مقایسه کنیم. یک طرح مربع ۵ × ۵ است. (۱۸۵) کتاب

عملتها (K) (col)

	۱	۲	۳	۴	۵	Y _{..}	تایارها	Y _{i..}
۱	A=11 B=5 C=4 D=1 E=1					-14	A	11
۲	B=-1 C=-1 D=5 E=2 A=11					9	B	-4
۳	C=-7 D=13 E=1 A=2 B=-4					5	C	-13
۴	D=1 E=4 A=1 B=-2 C=-3					3	D	4
۵	E=-3 A=5 B=-5 C=4 D=4					7	E	5
Y _{..k}	-11	11	-4	5	9	Y _{...}=10}		

داده های (row) (col)

شرایط میل - میل طرح - شرایط طرح - فرضیه ها نوشته شود.

$$SST_0 = \left[(-1)^2 + (-5)^2 + \dots + (4)^2 \right] - \frac{(10)^2}{25} = 474$$

$$SS_{Row} = \frac{[(-14)^2 + (9)^2 + (5)^2 + (3)^2 + (7)^2]}{5} - \frac{(10)^2}{25} = 41$$

$$SS_{col} = \frac{[(-11)^2 + (11)^2 + (-4)^2 + (5)^2 + (9)^2]}{5} - \frac{(10)^2}{25} = 150$$

$$SST = \frac{[11^2 + (-4)^2 + (-13)^2 + (4)^2 + (5)^2]}{5} - \frac{(10)^2}{25} = 330$$

$$PAPCOSSF = SST_0 - SST - SS_{Row} - SS_{col} = 121$$

source	df	SS	MS	F ₀
treatment	4	330	$MST = \frac{330}{4} = 82.5$	
Row	4	98	—	
col	4	150	—	$F_0 = \frac{MST}{MSE} = \frac{82.5}{10.147} = 8.12$
Error	12	128	$MSE = \frac{128}{12} = 10.67$	
Total	24	474		$P(F_{4,12}) = 3.18$ $F_0 > F_{\alpha} \Rightarrow R.H.$

نتیجه گیری آماری: با اطمینان ۹۵٪ تفاوت معنی داری بین تیمارها وجود دارد.

نتیجه گیری کلی: با اطمینان ۹۵٪ تفاوت معنی داری بین این ۵ نوع فصل دینامیک وجود دارد.

۱۴۰۳/۱۰/۱۱

طرح مربع یونانی لاتین: $P \times P$: ۱- یک طرح تک عاملی است یعنی هدف بررسی سطوح یک عامل (تیمارها) است.

۲- سه عامل محدود کننده در تصانیف کردن داریم: ۳- بقیه شرایط بسیار با اطمینان و تنها تیمارها متفاوت اند. ۴- تیمارها باید به تعداد زوجی

واحدهای آزمایشی اعمال شوند. ۵- مربع تک عاملی باشد. (کامل یعنی برای هر عامل محدود کننده همه تیمارها به کار رفته باشند.)

مربع یونانی لاتین: 3×3 نداریم از 4×4 شروع می شود.

هر چه تعداد عامل محدود کننده اضافه شود، به تعداد P افزایش می شود.

$$N = P^2$$

تعداد مشاهده در مربع یونانی لاتین: با مربع یونانی فرق می کند.

۳ عامل محدود کننده

- ۴ نوع خوراک
- ۴ نوع وضعیت جرمی
- ۴ راننده

۴ نوع لاتین

- A
- B
- C
- D

* حروف لاتین تیارها هستند *

حروف یونانی را برای عامل محدود کننده سوم نگار می‌بریم.

ستون	۱	۲	۳	۴	تیارها
۱	A α	B β	C γ	D δ	$i = 1, 2, \dots, p$
۲	B δ	A γ	D β	C α	سطر
۳	C β	D α	A δ	B γ	$K = 1, 2, \dots, p$
۴	D γ	C δ	B α	A β	ستون
					$J = 1, 2, \dots, p$
					یونانی
					$L = 1, 2, \dots, p$

در هر سطر و ستون همه تیارها باید به کار رفته باشند و از هر تیار یک بار نباید استفاده کرد. مثلاً A α نباید دوبار تکرار شود.

اندیس ۱ همیشه برای تیارها نگار می‌بریم (در کتاب جمله اینطوری نیستند)

اندیس ۲ برای ستون - اندیس ۳ برای سطر - اندیس ۴ برای یونانی

مدل طرح:

شرایط مدل:

$$V_{ijkl} = \mu + T_i + \beta_j + \gamma_k + \omega_l + \varepsilon_{ijkl}$$

هر چه عامل محدود کننده بیشتری کنیم، از خطاها پیاپی کم می‌شود.

$$\sum_i T_i = 0 \quad \sum_j \beta_j = 0 \quad \sum_k \gamma_k = 0 \quad \sum_l \omega_l = 0$$

هر یک از این اندیس‌ها باید ۰ باشند

$$\begin{cases} H_0: T_i = \dots = T_p = 0 \\ H_1: \text{Not } H_0 \end{cases}$$

فرضیه ۰

آنالیز واریانس ANOVA : تجزیه واریانس (کل به مولفه های خود در مدل)

$$SST_0 = SST + SS_{col} + SS_{row} + SS_u + SSE \quad N = p^2$$

$$df: N-1 = (p-1) + (p-1) + (p-1) + (p-1) + (p-1)(p-3)$$

$$SST_0 = ? SST + SS_{col} + SS_{row} + SS_u + SSE \quad \text{ترین:}$$

درجه آزادی (خطا از تقاضای لیبی) کل از بقیه نیست میار.

$$p^2 - 4(p-1) = (p-1)(p+1-4) = (p-1)(p-3) \quad \leftarrow \text{خطا} = \text{درجه آزادی}$$

* اگر فرض H: درسته حتماً باید مقایسات زوجی اینها هم *

اگر فرض H: پذیرفته شده بکار نیست مقایسات زوجی اینها هم.

$$SST_0 = \sum_i \sum_j \sum_k \sum_L y_{ijkl}^2 - \frac{y_{....}^2}{N}$$

$$SST = \sum_{i=1}^P \frac{y_{i...}^2}{P} - \frac{y_{....}^2}{N}$$

$$SS_{Row} = \sum_{k=1}^P \frac{y_{...k}^2}{P} - \frac{y_{....}^2}{N}$$

$$SS_{col} = \sum_{j=1}^P \frac{y_{.j..}^2}{P} - \frac{y_{....}^2}{N}$$

$$SS_u = \sum_{L=1}^P \frac{y_{...L}^2}{P} - \frac{y_{....}^2}{N}$$

$$SSE = SST_0 - SST - SS_{row} - SS_{col} - SS_u$$

جدول ANOVA :

Source	df	SS	MS	F ₀
treatment	p-1	SST	$MST = \frac{SST}{p-1}$	$F_0 = \frac{MST}{MSE} \stackrel{H_0}{=} F(p-1, (p-1)(p-3))$
col	p-1	SS _{col}		
row	p-1	SS _{row}		
u	p-1	SS _u		
Error	(p-1)(p-3)	SSE	$MSE = \frac{SSE}{(p-1)(p-3)}$	
Total	p ² -1	SST ₀		

$$R.H. : F_0 > F_{\alpha}(p-1, (p-1)(p-3))$$

تمرین : مثال ص ۱۹۵ و ص ۱۹۴ + ۱۰ و ۱۱ در ۲ ضرب کنیم. مثال ۵-۵

تمرین : برآورد پارامترهای مدل بارش و حاصل جواب.

تمرین : دو داده گشته $\leftarrow$ حتما از میانگین استفاده کنیم که نمره جدول را بدست می آوریم

نکته : کسانی که به روش $\min(SSE)$ حل می کنند، حتماً آفرید جواب نزنند باید جواب معقول بزنند

جوابی که معقول نیست حالتی داریم که نمی توانیم و حل کنیم : راهکار

دارد گشته : در طرح تصادفی ساده و یا حذف می کنیم یا میانگین را حذف می کنیم.

در طرح بلوکی : در صورت حذف داده گشته، طرح بلوکی ناقص می شود و از حالت کامل بودن خارج می شود و بهتر است را حذف شود.

انسانی با عطف اثبات داده کمیده و بهر مورد با استفاده از مدل می آید برای با اینهم
E(MS) ها از هر مدل (طرح) یکی حل کنیم و فقط با استفاده از E(MS) برای آزمون یک مقدار

- ۱- استفاده از میانگین (از هر بلوک - تبار کل) می توان استفاده کرد اما اگر بحث تبار
- فاست می توان یعنی بهتر است از میانگین تبارها استفاده کرد.
- ۲- عددی جای کمیده قرار دهیم نه SSE مدل کمترین باشد (بهینه ترین) ← بهتر
- است از این مورد استفاده کرد.

راه های دیگر

۱۴۰۳/۱۰/۳

برای یک داده کمیده در طرح بلوک

$$SSE = SSTo - SST - SSBlock = \sum_i \sum_j y_{ij}^2 - \frac{y_{..}^2}{N} - \sum_i \frac{y_{i.}^2}{b} + \frac{y_{..}^2}{N} - \sum_j \frac{y_{.j}^2}{a} + \frac{y_{..}^2}{N} = \sum_i \sum_j y_{ij}^2 - \sum_i \frac{y_{i.}^2}{b} - \sum_j \frac{y_{.j}^2}{a} + \frac{y_{..}^2}{N}$$

$$= y_{11}^2 + y_{12}^2 + \dots + y_{1b}^2 + y_{21}^2 + y_{22}^2 + \dots + y_{2b}^2 + \dots + y_{a1}^2 + y_{a2}^2 + \dots + y_{ab}^2 - \frac{y_{1.}^2 + \dots + (y_{i.}' + x)^2 + \dots + y_{a.}^2}{b}$$

$$= \frac{y_{1j}^2 + \dots + (y_{1j}' + x)^2 + \dots + y_{1b}^2}{a} + \frac{(y_{.j}' + x)^2}{ab}$$

$$\frac{\partial SSE}{\partial x} = y_{1j}^2 - \frac{2(y_{1j}' + x)}{b} - \frac{2(y_{.j}' + x)}{a} + \frac{2(y_{..}' + x)}{ab} = 0 \xrightarrow{\times ab} abx - ay_{1j}' - ax - by_{.j}' - bx$$

$$+ y_{..}' + x = 0 \Rightarrow x(ab - a - b + 1) = ay_{1j}' + by_{.j}' - y_{..}' \Rightarrow \hat{x} = \frac{ay_{1j}' + by_{.j}' - y_{..}'}{(a-1)(b-1)}$$

$y_{..}^2$ = جمع کل مشاهدات که تبار داده کمیده
 $y_{.j}^2$ = جمع بلوک (یا) که تبار داده کمیده
 $y_{i.}^2$ = جمع تبار (یا) که تبار داده کمیده

نکته: در طرح مربع لاتین و یونانی لاتین در صورت وجود داده گمشده همانند طرح بلوکی عمل می کنیم.

از درجه آزادی خطای باید به تعداد داده گمشده کم شود از درجه آزادی کل هم کم می شود. در جدول ANOVA هم باید از درجه آزادی کم شود.

اگر داده گمشده بود، یکی را x و دیگری را y در نظریه کلاسیک به x و دیگر به y منتقل می کنیم و معادله

۲ مجهول درست می آوریم.

بر آوردن یک داده گمشده در طرح مربع لاتین:

$$SSE = SST_0 - SST - SS_{Row} - SS_{Col} = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{..}^2}{N} - \sum_i \frac{y_{i.}^2}{P} + \frac{y_{..}^2}{N} - \sum_j \frac{y_{.j}^2}{P} + \frac{y_{..}^2}{N}$$

$$\sum_k \frac{y_{.k}^2}{P} + \frac{y_{..}^2}{N} = \sum_i \sum_j \sum_k y_{ijk}^2 - \sum_k \frac{y_{i.}^2}{P} - \sum_j \frac{y_{.j}^2}{P} - \sum_k \frac{y_{.k}^2}{P} + 2 \frac{y_{..}^2}{N} \Rightarrow SSE_n = y_{iii}^2 + \dots +$$

$$x^2 + \dots + y_{ppp}^2 - \frac{y_{i.}^2 + \dots + (y_{i.} + n)^2 + \dots + y_{p.}^2}{P} - \frac{y_{.j}^2 + \dots + (y_{.j} + n)^2 + \dots + y_{.p}^2}{P}$$

$$\frac{y_{..}^2 + \dots + (y_{..k} + n)^2 + \dots + y_{..p}^2}{P} + 2 \frac{(y_{..} + n)^2}{N}$$

$$\frac{\partial SSE_n}{\partial n} = 2n - \frac{2(y_{i.} + n)}{P} - \frac{2(y_{.j} + n)}{P} - \frac{2(y_{..k} + n)}{P} + \frac{2(y_{..} + n)}{P^2} = 0 \xrightarrow{\times \frac{P^2}{P^2}}$$

$$P^2 n - P(y_{i.} + n) - P(y_{.j} + n) - P(y_{..k} + n) + P(y_{..} + n) = 0 \Rightarrow n(P^2 - 3P + 2) = P y_{i.} + P y_{.j} +$$

$$P y_{..k} - 2 y_{..} \Rightarrow \hat{n} = \frac{P y_{i.} + P y_{.j} + P y_{..k} - 2 y_{..}}{(P-1)(P-2)}$$

برای درک داده کشیده (در طرح مربع یونانی لاتین):

$$SSE = SST_0 - SST - SS_{col} - SS_{row} - SS_u = \sum_i \sum_j \sum_u \sum_L y_{ijul}^2 - \frac{y_{...}^2}{N} - \sum_i \frac{y_{i...}^2}{P} + \frac{y_{...}^2}{N}$$

$$\sum_j \frac{y_{.j..}^2}{P} + \frac{y_{...}^2}{N} - \sum_u \frac{y_{...u.}^2}{P} + \frac{y_{...}^2}{N} - \sum_L \frac{y_{...L}^2}{P} + \frac{y_{...}^2}{N}$$

$$SSE_n = y_{1111}^2 + \dots + x^2 + \dots + y_{pppp}^2 - \frac{y_{i...}^2 + \dots + (y'_{i...} + x)^2 + \dots + y_{p...}^2}{P} - \frac{y_{.j..}^2 + \dots + (y'_{.j..} + x)^2 + \dots + y_{.p.}^2}{P}$$

$$- \frac{y_{...u.}^2 + \dots + (y'_{...u.} + x)^2 + \dots + y_{...p}^2}{P} - \frac{y_{...L}^2 + \dots + (y'_{...L} + x)^2 + \dots + y_{...p}^2}{P} + \frac{3(y'_{...} + x)^2}{P^2 = N}$$

$$\frac{2(y'_{...L} + x)}{P} + \frac{4(y'_{...} + x)}{P^2} \xrightarrow{x \cdot \frac{P^2}{P}} p^2 x - p y'_{i...} - p x - p y'_{.j..} - p x - p y'_{...k.} - p x - p y'_{...L}$$

$p^2 - (p + 3) = (p-1)(p-3)$

$$- p x + 3 y'_{...} + 3 x = 0 \Rightarrow (p^2 - p - p - p - p + 3) x = p(y'_{i...} + y'_{.j..} + y'_{...k.} + y'_{...L}) - 3 y'_{...}$$

$$\hat{x} = \frac{p(y'_{i...} + y'_{.j..} + y'_{...u.} + y'_{...L}) - 3 y'_{...}}{(p-1)(p-3)}$$

آنگاه داده کشیده مربوط به تیمار A بود و در هر A ها را جمع کنیم بجز A مربوط به داده کشیده.

آزمون LSD مقایسه زوجی در مربع یونانی لاتین:

$$LSD = t_{\frac{\alpha}{2}} (df_E) \cdot \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$$

در اینجا n_i ها همان p هستند $\frac{1}{n_i} \Rightarrow \frac{1}{p}$

درجه آزادی (dfE):
 - مربع: $(p-1)(p-3)$
 - یونانی: $(p-1)(p-3)$