IST3002 Deney Tasarımı

Latin Kare Tasarım

Fatih Kızılaslan

Marmara Üniversitesi

2019-2020 Bahar X. Hafta

Latin Kare Tasarım

- Rastgele blok tasarımında bir yönlü ANOVA'da olduğu gibi etkisi araştırılmak istenen bir ana faktör ile bir tane bloklama faktörü kullanılır.
- Bloklar kendi içinde **homojen** ve bloklar kendi aralarında **heterojen** olarak oluşturulur.
- Deney birimleri arasındaki sistematik farklılıkların etkisini gidermek amacıyla bloklama yapılır [**Senoglu ve Acitas**].
- Eğer deney birimleri arasındaki heterojenlik bir tane bloklama faktörü ile giderilemeyecek kadar fazla ise, iki farklı bloklama faktörü kullanılarak homojenlik sağlanmaya çalışılır. Böylece, deneysel hata azaltılmış olur [**Senoglu ve Acitas**]. Bu tip tasarımlar **Latin kare tasarım** (Latin square design) olarak adlandırılır.
- Latin kare tasarımında, **bir ana faktör ve iki bloklama faktörü** kullanılır.

Birdal Şenoğlu ve Şükrü Acıtaş (2014). İstatistiksel Deney Tasarımı Sabit Etkili Modeller, 3. Basım, Nobel Akademik Yayınevi.

- Latin kare tasarımda satır sayısı, sütün sayısı ve deneme (ana faktörün düzeyleri) sayısı **birbirine eşit** olmalıdır. **(1. Kısıt)**
- Latin kare tasarımda satır ve sütunlarda bloklama faktörleri kullanılır. Her faktör düzeyi (deneme) Latin harfleri ile her satır ve her sütunda **yalnızca bir kez** gözlenir. **(2. Kısıt)**
- Latin kare tasarımda rastgelelik üzerine 2 kısıt vardır. Kolay bir tasarım olmasına rağmen bu kısıtlar nedeniyle çok tercih edilmez.
- Latin karesinin boyutlarını pxp **Latin karesi** biçiminde ifade ederiz.
- p faktörlü bir Latin karesi veya pxp Latin karesi p satır ve p sütundan oluşan bir karedir. Toplam gözlem sayısı p^2 dir.

Örneğin, bir faktörümüzün A, B ve C gibi üç düzeyi olsun. Bu durumda 3x3 tipinde 12 Latin kare tasarımı vardır. Latin kare tasarımda karenin her bir satır ve sütununda üç faktör düzeyi A, B ve C'nin tam bir tekrarı vardır.

$$\text{Satırlar}\left\{\begin{array}{ccc} \multicolumn{3}{c}{\overbrace{\hphantom{A \quad B \quad C}}^{\text{Sütunlar}}} \\ A & B & C \\ B & C & A \\ C & A & B \end{array}\right. , \quad \begin{array}{ccc} B & A & C \\ C & B & A \\ A & C & B \end{array} , \quad \begin{array}{ccc} C & B & A \\ B & A & C \\ A & C & B \end{array}$$

Latin Kare Tasarımı: Matematiksel Model

pxp Latin kare tasarımı için matematiksel model

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \epsilon_{ijk}, \ \ i,j,k = 1,...,p \tag{1}$$

biçiminde ifade edilir. Burada,

y_{ijk} : i. satır, k. sütunda j. düzeye (denemeye) ait olan gözlem değerini,
μ : genel ortalamayı,
α_i : i. satırın etkisini,
τ_j : faktörün j. düzeyinin etkisini,
β_k : k. sütunun etkisini,
ϵ_{ijk} : rastgele hata terimini

gösterir. **Varsayım:** $\epsilon_{ijk} \sim N(0, \sigma^2)$, $i,j,k = 1,...,p$ biribirinden bağımsız rastgele değişkenlerdir

(1) Latin kare modelinin **sabit etkili** olduğunu varsayalım (yani tüm faktör düzeyleri özel olarak seçilmiştir). Bu durumda

$$\sum_{i=1}^{p} \alpha_i = 0, \ \sum_{j=1}^{p} \tau_j = 0 \text{ ve} \sum_{k=1}^{p} \beta_l = 0$$

olur.

Latin kare tasarımında hipotezlerimiz aşağıdaki gibi olur. Blok faktörler (satır ve sütun) için testler genellikle tercih edilmez.

- Ana faktörün düzeyleri için $H_0 : \tau_1 = \tau_2 = ... = \tau_p = 0$ ve H_1 :En az bir i için $\tau_i \neq 0$
- Satır bloklama faktörü için $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_p = 0$ ve H_1 :En az bir i için $\alpha_i \neq 0$
- Sütun bloklama faktörü için $H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0$ ve H_1 :En az bir i için $\beta_i \neq 0$

Kareler Toplamının Parçalanışı

Diğer ANOVA modellerinde olduğu gibi burada da kareler toplamları

$$
\begin{aligned}
SS_T &= \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}(y_{ijk}-\overline{y}_{...})^2\,,\; SS_{Deneme} = p\sum_{i=1}^{p}(\overline{y}_{i..}-\overline{y}_{...})^2 \\
SS_{Satır} &= p\sum_{i=1}^{p}\left(\overline{y}_{.j.}-\overline{y}_{...}\right)^2,\; SS_{Sutun} = p\sum_{i=1}^{p}(\overline{y}_{..k}-\overline{y}_{...})^2 \text{ ve} \\
SS_E &= \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}\left(y_{ijk}-\overline{y}_{i..}-\overline{y}_{.j.}-\overline{y}_{..k}+2\overline{y}_{...}\right)^2
\end{aligned}
$$

olmak üzere

$$SS_T = SS_{Deneme} + SS_{Satır} + SS_{Sutun} + SS_E$$

biçiminde bileşenlerine ayrılır.

Test İstatistikleri

Latin kare modeli için yukarıda verdiğimiz hipotezler aşağıdaki gibi sınarız.

- Eğer $F_{Deneme} = \frac{SS_{Deneme}/(p-1)}{SS_E/(p-2)(p-1)} = \frac{MS_{Deneme}}{MS_E}$ test istatistiğinin değeri $F_{Deneme} > F_{(p-1),(p-2)(p-1),\alpha}$ olur ise

$$H_0 : \tau_1 = \tau_2 = ... = \tau_p = 0$$

hipotezi reddedilir.

- Eğer $F_{Satır} = \frac{SS_{Satır}/(p-1)}{SS_E/(p-2)(p-1)} = \frac{MS_{Satır}}{MS_E}$ test istatistiğinin değeri $F_{Satır} > F_{(p-1),(p-2)(p-1),\alpha}$ olur ise

$$H_0 : \alpha_1 = \alpha_2 = ... = \alpha_p = 0$$

hipotezi reddedilir.

- Eğer $F_{Sutun} = \frac{SS_{Sutun}/(p-1)}{SS_E/(p-2)(p-1)} = \frac{MS_{Sutun}}{MS_E}$ test istatistiğinin değeri $F_{Sutun} > F_{(p-1),(p-2)(p-1),\alpha}$ olur ise

$$H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0$$

ANOVA Tablosu

Latin kare tasarımı için ANOVA tablosu aşağıdaki gibi oluşturulur.

Değişim kaynağı	Kareler toplamı	Serbestlik derecesi	Kareler ortalaması	F test değeri
Denemeler	SS_{Deneme}	$p-1$	$MS_{Deneme} = \frac{SS_{Deneme}}{p-1}$	$F_{Deneme} = \frac{MS_{Deneme}}{MS_E}$
Satırlar	$SS_{Satır}$	$p-1$	$MS_{Satır} = \frac{SS_{Satır}}{p-1}$	$F_{Satır} = \frac{MS_{Satır}}{MS_E}$
Sütunlar	SS_{Sutun}	$p-1$	$MS_{Sutun} = \frac{SS_{Sutun}}{p-1}$	$F_{Sutun} = \frac{MS_{Sutun}}{MS_E}$
Hata	SS_E	$(p-2)(p-1)$	$MS_E = \frac{SS_E}{(p-2)(p-1)}$	
Toplam	SS_T	p^2-1		

Not: $(p^2-1)-(p-1)-(p-1)-(p-1)=(p-2)(p-1)$ dir. Latin kare tasarımında p

Parametre Tahmini

Latin kare tasarım modelinde (1) bilinmeyen parametreler $\mu, \alpha_i, \tau_j, \beta_k$, $i, j, k = 1, ..., p$ için *en küçük kareler (EKK)* tahmin edicileri

$$\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}\epsilon_{ijk}^2 = \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p}(y_{ijk} - \mu - \alpha_i - \tau_j - \beta_k)^2$$

kareler toplamını minimize eden değerleri bulunarak belirlenir. Bu durumda,

$$\begin{aligned}
\widehat{\mu} &= \overline{y}_{...} \\
\widehat{\alpha}_i &= \overline{y}_{i..} - \overline{y}_{...}, \ i = 1, ..., p \\
\widehat{\tau}_j &= \overline{y}_{.j.} - \overline{y}_{...}, \ j = 1, ..., p \\
\widehat{\beta}_k &= \overline{y}_{..k} - \overline{y}_{...}, \ k = 1, ..., p
\end{aligned} \tag{2}$$

biçiminde elde edilir.

Böylece, (1) modeli için yanıt değişkeni y_{ij}' nin tahmin edicisi

$$
\begin{aligned}
\widehat{y}_{ijk} &= \widehat{\mu} + \widehat{\alpha}_i + \widehat{\tau}_j + \widehat{\beta}_k \\
&= \overline{y}_{...} + (\overline{y}_{i..} - \overline{y}_{...}) + (\overline{y}_{i..} - \overline{y}_{...}) + (\overline{y}_{..k} - \overline{y}_{...}) \\
&= \overline{y}_{i..} + \overline{y}_{i..} + \overline{y}_{..k} - 2\overline{y}_{...} \qquad (3)
\end{aligned}
$$

olur.

Böylece, modeldeki artıklar e_{ijk}, $i, j, k = 1, ..., p$ olmak üzere

$$
\begin{aligned}
e_{ijk} &= y_{ijk} - \widehat{y}_{ijk} \\
&= y_{ijk} - \overline{y}_{i..} - \overline{y}_{.j.} - \overline{y}_{..k} + 2\overline{y}_{...} \qquad (4)
\end{aligned}
$$

biçiminde bulunur.
Artıkları kullanılarak **modelin varsayımlarını kontrol ederiz.**

Eksik (Kayıp) Gözlem Durumu

pxp Latin kare tasarımı uygulanan bir deneyde herhangi bir $y_{ijk} = x$ gözleminin eksik olması durumunda

$$\widehat{x} = \frac{p(\overline{y}^*_{i..} + \overline{y}^*_{i..} + \overline{y}^*_{..k}) - 2\overline{y}^*_{...}}{(p-2)(p-1)}$$

biçiminde tahmin edilir.

- Eksik gözlem $y_{ij} = x$ yerine EKK tahmin edicisi $\widehat{x}$ yazılarak kareler toplamları ve ANOVA tablousu oluşturulur.
- Her bir eksik gözlem için hatanın serbestlik derecesi 1 azaltılır.

pxp Latin kare tasarımında ANOVA tablosunda hatanın serbestlik derecesi p'ye bağlıdır. p küçük olduğunda serbestlik derecesi de küçük olacaktır. Bu durumda hesaplacak olan F tablo değerleri de büyük olacaktır. Hatanın serbestlik derecesini büyütmek için tekrar sayısının arttırılması önerilir. Birkaç farklı yöntem ile tekrar sayısı arttırılabilir.

Örnek 1: Bir fabrikada çalışan bir mühendis 5 farklı ekip (A, B, C, D, E) tarafından üretilen ürün miktarları ile ilgili istatistiksel analiz yapmak istiyor. Üretilen ürün miktarında kullanılan 5 farklı malzeme türünün ve mesai gününün (hafta içi 5 gün) etkisi olduğu düşünülmektedir. 5 farklı ekibin 5 farklı malzeme ve 5 farklı günde elde ettiği ürün miktarları aşağıdaki gibi elde edilmiştir.

	Günler				
Malzeme	1	2	3	4	5
I	$A = 8$	$B = 7$	$D = 1$	$C = 7$	$E = 3$
II	$C = 11$	$E = 2$	$A = 7$	$D = 3$	$B = 8$
III	$B = 4$	$A = 9$	$C = 10$	$E = 1$	$D = 5$
IV	$D = 6$	$C = 8$	$E = 6$	$B = 6$	$A = 10$
V	$E = 4$	$D = 2$	$B = 3$	$A = 8$	$C = 8$

$\alpha = 0.05$ anlamlılık düzeyinde bu veriyi kullanarak üretim miktarındaki farklılıkları istatistiksel olarak açıklayınız.