IST3002 Deney Tasarımı

Rastgele Blok Tasarımı: Etkileşimli Model ve Rastgele Etkili Model

Fatih Kızılaslan

Marmara Üniversitesi

2019-2020 Bahar IX. Hafta

Rastgele Blok Tasarımı: Birden Fazla Gözlem Durumu

a faktör düzeyine sahip bir faktör, b farklı blok, her bir blokta her bir düzeyin n defa tekrarlandığı ve faktör düzeyleri ile bloklar arasında **etkileşimin olmadığı** bir rastgele tam blok tasarımı için matematiksel model

$$y_{ijk} = \mu + \tau_i + \beta_j + \epsilon_{ijk}, \;\; i = 1, ..., a, \;\; j = 1, ..., b, \;\; k = 1, ..., n \qquad (1)$$

biçiminde ifade edilir.

Bu model için veri yapısı aşağıdaki gibi olur.

Düzeyler (Denemeler)	**Bloklar**												
	1. Blok				2. Blok				. . .	b. Blok			
1	y_{111},	y_{112},	$\ldots$	y_{11n}	y_{121},	y_{122},	$\ldots$,	y_{12n}	. . .	y_{1b1},	y_{1b2},	$\ldots$,	y_{1bn}
2	y_{111},	y_{112},	$\ldots$	y_{11n}	y_{221},	y_{222},	$\ldots$,	y_{22n}	. . .	y_{2b1},	y_{2b2},	$\ldots$,	y_{2bn}
.	.	.	.	.	.	.	.	.	. . .	.	.	.	.
.	.	.	.	.	.	.	.	.	. . .	.	.	.	.
.	.	.	.	.	.	.	.	.	. . .	.	.	.	.
a	y_{a11},	y_{a12},	$\ldots$	y_{a1n}	y_{a21},	y_{a22},	$\ldots$,	y_{a2n}	. .	y_{ab1},	y_{ab2},	$\ldots$,	y_{abn}

Bu durumda, toplam gözlem sayısı $N = abn$ ve

- $y_{i..} = \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}$, $i = 1, ..., a$ ve $\overline{y}_{i..} = \frac{y_{i..}}{b\ n} \rightarrow i.$ faktör düzeyindeki gözlemler (satırlar) için ortalama
- $y_{.j.} = \sum_{i=1}^{a} \sum_{k=1}^{n} y_{ijk}$, $j = 1, ..., b$ ve $\overline{y}_{.j.} = \frac{y_{.j.}}{a\ n} \rightarrow j.$ bloktaki gözlemler (sütunlar) için ortalama
- $y_{...} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}$ ve $\overline{y}_{...} = \frac{y_{...}}{N} \rightarrow$ tüm gözlemler için genel ortalama

biçiminde olur.

Ayrıca,

$$
\begin{aligned}
SS_T &= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \overline{y}_{...})^2 \\
SS_{Deneme} &= bn \sum_{i=1}^{a} (\overline{y}_{i..} - \overline{y}_{...})^2, \ SS_{Blok} = an \sum_{j=1}^{b} \left(\overline{y}_{.j.} - \overline{y}_{...}\right)^2 \\
SS_E &= \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \left(y_{ijk} - \overline{y}_{i..} - \overline{y}_{.j.} + \overline{y}_{...}\right)^2
\end{aligned}
$$

olmak üzere $SS_T = SS_{Deneme} + SS_{Blok} + SS_E$ biçiminde parçalanır.

ANOVA Tablosu

Hipotezler ve tüm analiz her blokta her düzey için **sadece 1 gözlemin** olduğu durum ile aynı biçimde yapılır. Hipotezlerimiz ve ANOVA tablosu aynı biçimde olacaktır. Sadece $N = abn$ omak üzere toplamlarda bazı değişiklikler olur.

Değişim kaynağı	Kareler toplamı	Serbestlik derecesi	Kareler ortalaması	F test değeri
Denemeler	SS_{Deneme}	$a-1$	$MS_{Deneme} = \frac{SS_{Deneme}}{a-1}$	$F_{Deneme} = \frac{MS_{Deneme}}{MS_E}$
Bloklar	SS_{Blok}	$b-1$	$MS_{Blok} = \frac{SS_{Blok}}{b-1}$	$F_{Blok} = \frac{MS_{Blok}}{MS_E}$
Hata	SS_E	$\underbrace{N-a-b+1}_{(a-1)(b-1)}$	$MS_E = \frac{SS_E}{(a-1)(b-1)}$	
Toplam	SS_T	$N-1$		

Rastgele Blok Tasarımı: Birden Fazla Gözlem ve Etkileşim Durumu

a faktör düzeyine sahip bir faktör, b farklı blok, her bir blokta her bir düzeyin n defa tekrarlandığı ve faktör düzeyleri ile bloklar arasında **etkileşimin** olduğu bir rastgele tam blok tasarımı için matematiksel model

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}, \;\; i = 1, ..., a, \;\; j = 1, ..., b, \;\; k = 1, ..., n \quad (2)$$

biçiminde ifade edilir.
$(\tau\beta)_{ij}$: i. faktör düzeyi ile j. blok arasındaki etkileşimin etkisini gösterir.

(2) modeli sabit etkili model olduğunda

$$\sum_{i=1}^{a} \tau_i = 0, \;\; \sum_{j=1}^{b} \beta_j = 0 \text{ ve } \sum_{i=1}^{a} (\tau\beta)_{ij} = \sum_{j=1}^{b} (\tau\beta)_{ij} = 0$$

olur.

Hipotezler

Sabit etkili (2) modeli için aşağıdaki hipotezleri test edebiliriz.

- Faktör düzeylerinin etkisi (ortalamaları) için

$$\begin{aligned} H_0 &: \tau_1 = \tau_2 = \ldots = \tau_a = 0 \\ H_1 &: \text{En az bir } i \text{ için } \tau_i \neq 0 \end{aligned} \tag{3}$$

- Blokların etkisi için

$$\begin{aligned} H_0 &: \beta_1 = \beta_2 = \ldots = \beta_b = 0 \\ H_1 &: \text{En az bir } j \text{ için } \beta_j \neq 0 \end{aligned} \tag{4}$$

- Faktör düzeyleri ile blokların etkileşimlerinin etkisi için

$$\begin{aligned} H_0 &: (\tau\beta)_{11} = (\tau\beta)_{12} = \ldots = (\tau\beta)_{ab} = 0 \\ H_1 &: \text{En az bir } (i,j) \text{ için } (\tau\beta)_{ij} \neq 0 \end{aligned} \tag{5}$$

Kareler Toplamının Parçalanışı

(2) modeli için kareler toplamları

$$
\begin{aligned}
SS_T &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\overline{y}_{...})^2 \\
SS_{Deneme} &= bn\sum_{i=1}^{a}(\overline{y}_{i..}-\overline{y}_{...})^2\,,\ SS_{Blok} = an\sum_{j=1}^{b}\left(\overline{y}_{.j.}-\overline{y}_{...}\right)^2 \\
SS_{Etkilesim} &= n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{y}_{ij.}-\overline{y}_{i..}-\overline{y}_{.j.}+\overline{y}_{...}\right)^2 \\
SS_E &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\overline{y}_{ij.}\right)^2
\end{aligned}
$$

olmak üzere $SS_T = SS_{Deneme} + SS_{Blok} + SS_{Etkilesim} + SS_E$ eşitliği geçerlidir.

Test İstatistikleri ve Kurallar

1. Eğer $F_{Deneme} = \frac{MS_{Deneme}}{MS_E} > F_{(a-1),N-ab,\alpha}$ olur ise

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

   hipotezi reddedilir. Bu durumda, *faktör düzeyleri arasında anlamlı bir farklılık vardır.*

2. Eğer $F_{Blok} = \frac{MS_{Blok}}{MS_E} > F_{(b-1),N-ab,\alpha}$ olur ise

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

   hipotezi reddedilir. Bu durumda, *bloklar arasında anlamlı bir farklılık vardır.*

3. Eğer $F_{Etkilesim} = \frac{MS_{Etkilesim}}{MS_E} > F_{(a-1)(b-1),N-ab,\alpha}$ olur ise

   $$H_0 : (\tau\beta)_{11} = (\tau\beta)_{12} = ... = (\tau\beta)_{ab} = 0$$

   hipotezi reddedilir. Bu durumda, *faktör düzeyleri ile bloklar arasındaki etkileşim anlamlıdır.*

ANOVA Tablosu

Değişim kaynağı	Kareler toplamı	Serbestlik derecesi	Kareler ortalaması	F test değeri
Denemeler	SS_{Deneme}	$a-1$	$MS_{Deneme} = \dfrac{SS_{Deneme}}{a-1}$	$F_{Deneme} = \dfrac{MS_{Deneme}}{MS_E}$
Bloklar	SS_{Blok}	$b-1$	$MS_{Blok} = \dfrac{SS_{Blok}}{b-1}$	$F_{Blok} = \dfrac{MS_{Blok}}{MS_E}$
Etkileşim	$SS_{Etkilesim}$	$(a-1)(b-1)$	$MS_{Etkilesim} = \dfrac{SS_{Etkilesim}}{(a-1)(b-1)}$	$F_{Etkilesim} = \dfrac{MS_{Etkilesim}}{MS_E}$
Hata	SS_E	$N-ab$	$MS_E = \dfrac{SS_E}{(N-ab)}$	
Toplam	SS_T	$N-1$		

Not: Burada $N = abn$ olmak üzere
$N-1-(a-1)-(b-1)-(a-1)(b-1) = N-ab$ dir.

Rastgele Blok Tasarımı: Rastgele Etkili (Random Effects) Model

- Bir rastgele tam blok tasarımında faktör düzeyleri, ve/veya bloklar ve/veya faktör düzeyi ile blokların etkileşimleri rastgele etkili olabilirler.
- Eğer bir deneyde kullanılan faktörlerin bazılarının düzeyleri rastgele seçiliyorsa (rastgele etkili) ve diğer faktörlerin düzeyleri özel olarak seçiliyorsa (sabit etkili) bağımlı (yanıt) değişkeni ile bu faktörlerin düzeyleri arasındaki ilişkiye veren modele **karışık etkili model** (mixed effects model) denir.
- Şimdiye kadar faktör düzeylerinin ve blokların özel olarak seçilmiş yani sabit etkili olduğu varsayımını yaptık.
- Faktör düzeylerinin veya blokların veya her ikisininde rastgele etkili olabileceği birçok durum vardır.
- Şimdi, faktör düzeylerinin **sabit etkili** ve blokların **rastgele etkili** olduğu modeli inceleyelim.

a faktör düzeyi **özel olarak** seçilen bir faktör, *b* tane bloğun **rastgele** seçildiği ve her bir blokta her bir düzeyin 1 defa tekrarlandığı bir rastgele tam blok tasarımı için matematiksel model

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \;\; i = 1, ..., a, \;\; j = 1, ..., b, \qquad (6)$$

biçimindedir.

- (6) modeli karışık etkili bir modeldir.
- Faktör düzeyleri sabit etkili olduğundan $\sum_{i=1}^{a} \tau_i = 0$ olur.
- Bloklar rastgele etkili olduğundan $\beta_j \sim N(0, \sigma_\beta^2)$, $j = 1, ..., b$ birbirinden ve ϵ_{ij}'den bağımsız rastgele değişkenler olduğu varsayımı yapılır.

(6) modeli için

$$
\begin{aligned}
E(y_{ij}) &= \mu + \tau_i, i = 1, ..., a \\
Var(y_{ij}) &= Var(\beta_j) + Var(\epsilon_{ij}) = \sigma^2_\beta + \sigma^2 \\
Cov(y_{ij}, y_{ij^*}) &= 0,\ j \neq j^* \text{ için (farklı bloklardaki gözlemlerin kovaryansı)} \\
Cov(y_{ij}, y_{i^*j}) &= \sigma^2_\beta,\ i \neq i^* \text{ için (aynı bloklardaki gözlemlerin kovaryansı)}
\end{aligned}
$$

biçiminde bulunur.

Hipotezlerimiz:

- Faktör düzeylerinin etkisi için: $H_0 : \tau_1 = \tau_2 = ... = \tau_a = 0$ ve H_1 :En az bir i için $\tau_i \neq 0$
- Blokların etkisi için: $H_0 : \sigma^2_\beta = 0$ ve $H_1 : \sigma^2_\beta > 0$

biçiminde olur. (Araştırmacının isteğine bağlı olarak yapılır.)

1. Eğer $F_{Deneme} = \frac{MS_{Deneme}}{MS_E} > F_{(a-1),N-ab,\alpha}$ olur ise

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

hipotezi reddedilir. Bu durumda, *faktör düzeyleri arasında anlamlı bir farklılık vardır.*

2. Eğer $F_{Blok} = \frac{MS_{Blok}}{MS_E} > F_{(b-1),N-ab,\alpha}$ olur ise

$$H_0 : \sigma_\beta^2 = 0$$

hipotezi reddedilir. Bu durumda, *bloklar arasında (değişkenlik) anlamlı bir farklılık vardır.*

Ayrca, varyans bileşenleri σ^2 ve σ_β^2 için tahmin ediciler

$$\widehat{\sigma}^2 = MS_E \text{ ve } \widehat{\sigma}_\beta^2 = \frac{MS_{Blok} - MS_E}{a}$$

biçimindedir.

Etikleşimli Karışık Etkili Model

Eğer (6) modelinde faktör düzeyleri ile bloklar arasında etkileşim var ise, bu durumda modelimiz aşağıdaki gibi olur.

a faktör düzeyi **özel olarak** seçilen bir faktör, b tane bloğun **rastgele** seçildiği, her bir blokta her bir düzeyin 1 defa tekrarlandığı ve faktör düzeyleri ile bloklar arasında **etkileşimin** olduğu bir rastgele tam blok tasarımı için matematiksel model

$$y_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ij}, \;\; i = 1, ..., a, \;\; j = 1, ..., b, \tag{7}$$

biçiminde ifade edilir.

- Faktör düzeyleri sabit etkili olduğundan $\sum_{i=1}^{a} \tau_i = 0$ olur.
- **Varsayım:** Bloklar rastgele etkili oluğundan $\beta_j \sim N(0, \sigma_\beta^2)$ birbirinden ve ϵ_{ij}'den bağımsızdır
- Ayrıca, etkileşim terimi rassal olan blokları içerdiğinde $(\tau\beta)_{ij}$ de rastgele değişken olur. **Varsayım:** $(\tau\beta)_{ij} \sim N(0, \sigma^2_{\tau\beta})$, $i = 1, ..., a, j = 1, ..., b$ birbirinden ve ϵ_{ij}'den bağımsızdır.

Faktör düzeylerinin etkisini test etmek için test istatistiğimiz ve kuralımız aynıdır.

Eğer $F_{Deneme} = \frac{MS_{Deneme}}{MS_E} > F_{(a-1),N-ab,\alpha}$ olur ise

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

hipotezi reddedilir. Bu durumda, *faktör düzeyleri arasında anlamlı bir farklılık vardır.*

Eksik (Kayıp) Gözlem Durumu

Rastgele etkili model tasarımını kullanarak varyans analizi yaparken bazı durumlarda eksik gözlemler ile karşılaşabiliriz.

a faktör düzeyine sahip olan bir faktör ve b farklı bloğun olduğu bir rastgele tam blok tasarımında

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}, \;\; i = 1, ..., a, \;\; j = 1, ..., b \tag{8}$$

j. blokdaki i. düzey için gözlem y_{ij}'nin kayıp (eksik) olduğunu varsayalım.

Bu durumda veri yapısı aşağıdaki gibi olur.

Düzeyler (Denemeler)	**Bloklar** 1	2	.	j	.	b	**Toplam**
1	y_{11}	y_{12}	.	.	.	y_{1b}	$y_{1.}$
.	.	.	.	.	.		.
i	y_{i1}	y_{i2}.	.	x (Kayıp)	.		$y_{i.}^* + x$
.	.	.	.	.	.		.
a	y_{a1}	y_{a2}	.	.	.	y_{ab}	$y_{a.}$
Toplam	$y_{.1}$	$y_{.2}$	.	$y_{.j}^* + x$	.	$y_{.b}$	$y_{..}^* + x$

$y_{i.}^*$: i. faktör düzeyinde eksik gözlem x dışındaki diğer gözlemlerin toplamı
$y_{.j.}^*$: j. blokta eksik gözlem x dışındaki diğer gözlemlerin toplamı

- x eksik gözlemi için en küçük kareler (EKK) tahmin edicisi bulalım. Amacımız (8) modelinde hata kareler toplamı SS_E'yi minimum yapacak x değerini belirlemektir.
-
$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\left(y_{ij} - \overline{y}_{i.} - \overline{y}_{.j} + \overline{y}_{..}\right)^2$$

  toplamını x'e göre minimum yapan değeri $\frac{d(SS_E)}{dx} = 0$ denklemini çözerek

$$\widehat{x} = \frac{ay_{i.}^* + by_{.j}^* - y_{..}^*}{(a-1)(b-1)} \tag{9}$$

  biçiminde elde ederiz.

Eksik Gözlem ile Analiz

- Eksik gözlem $y_{ij} = x$ yerine EKK tahmin edicisi (9)'da verilen $\widehat{x}$ yazılarak kareler toplamları ve ANOVA tablousu oluşturulur.
- Her bir eksik gözlem için hatanın serbestlik derecesi 1 azaltılır.

Örneğin, sadece $y_{ij} = x$ gözlemi eksik olduğunda ANOVA tablosu aşağıdaki gibi olacaktır.

Değişim kaynağı	Kareler toplamı	Serbestlik derecesi	Kareler ortalaması	F test değeri
Denemeler	SS_{Deneme}	$a-1$	MS_{Deneme}	F_{Deneme}
Bloklar	SS_{Blok}	$b-1$	MS_{Blok}	F_{Blok}
Hata	SS_E	$N-a-b$	MS_E	
Toplam	SS_T	$N-2$		