BÖLÜM III

RASTGELE BLOK DÜZENİ VE GENELLEŞTİRİLMİŞ RASTGELE BLOK DÜZENİ (İKİ YÖNLÜ VARYANS ANALİZİ)

III.1 RASTGELE BLOK DÜZENİ

Deney düzeninde gözlemler arasındaki farklılıklar, hata varyansına önemli etki yapabilmektedirler. Ancak; bu etkiler farklı çevre koşulları (bir günün farklı zamanları, yılın mevsimleri, farklı günler, farklı sınıflar v.s.) nedeniyle görülmeyebilir. Bu tür etkileri ortaya çıkarabilmek ve böylece hata varyansını küçültebilmek için bloklama yapmaya ihtiyaç vardır. Çünkü, bağımlı değişkene ait değişimin (varyansın) bir kaynağını oluşturan ve gürültü değişkeni olarak adlandırılan görülemeyen bu tür etkiler kontrol altında tutulmalıdır. Bu istenmeyen değişkeni kontrol altında tutabilmenin en basit yolu, gürültü değişkenini sabit tutmaktır. Yani, deney aynı gün, aynı saat, aynı yaş grubunda yapılacak şekilde planlanmalıdır. Eğer bunu yapma imkanı yoksa, gürültü değişkenini bir başka faktör olarak ele alıp modele katmak gerekir. Böylece gürültü değişkeni, bloklama değişkeni olarak kabul edilir ve N tane gözlem (örnek birimi) n tane bloğa rastgele olarak dağıtılır. Ana faktör veya deneme düzeylerinin sayısı k olmak üzere, kısıtlayıcısız tek faktör rastgele deney düzenine bloklama değişkeni katılarak rastgelelik üzerine bir kısıtlama getirilmiş olmaktadır. Çünkü bloklar oluşturulurken, bir bloktaki gözlemler (birimler) diğer bloklardaki gözlemlere göre daha homojen olmalıdır. Örneğin;

Faktör (Deneme):Tedavi Metodu…… Düzey sayısı $k = 4$ $(M1, M2, M3, M4)$

Blok: Yaş grubu…. Blok Sayısı $n = 3$, $(B1, B2, B3)$

Cözlem (deney) sayısı $N = k * n = 12$

Deney Düzeni: Her blok, her denemede bir defa görünür, birimler bloklar içinde bağımsız, rastgele ve homojen iken bloklar arasında heterojen olacak şekilde rastgelelik üzerine blok kısıtı getirilmektedir.

<table>
<tr><th rowspan="2">BLOK</th><th colspan="4">TEDAVİ METODU</th><th rowspan="2"></th><th rowspan="2"></th></tr>
<tr><th>M1</th><th>M2</th><th>M3</th><th>M4</th></tr>
<tr><td>B1(20-24)</td><td>G1</td><td>G2</td><td>G3</td><td>G4</td><td>Gözlemler kendi içinde bağımsız ve rastgele, homojen</td><td rowspan="3">Gözlemler bloklar arasında heterojen</td></tr>
<tr><td>B2(25-29)</td><td>G5</td><td>G6</td><td>G7</td><td>G8</td><td>Gözlemler kendi içinde bağımsız ve rastgele, homojen</td></tr>
<tr><td>B3(30-34)</td><td>G9</td><td>G10</td><td>G11</td><td>G12</td><td>Gözlemler kendi içinde bağımsız ve rastgele, homojen</td></tr>
</table>

Rastgele blok düzeninin varsayımları:

1. Rastgele blok düzeni modeli, bağımlı değişkende meydana gelen değişimi etkileyen tüm kaynakları (deneme, blok ve hata) kapsar.

2. Deney, tüm denemelerden sadece ilgilenilenleri kapsar.

3. Denemelerin sayısı $k \geq 2$ ve blok değişkeninin düzey sayısı $n \geq 2$ olmalıdır.

4. Her bir blokta, k tane homojen gözlem yer almalıdır.

5. Denemeler her bir blok içinde yer alan gözlem birimlerine rastgele atanmalıdır.

6. Hata terimi $\varepsilon_{ij} \sim BND(0\ ,\ \sigma_\varepsilon^2)$ olmalıdır.

Rastgele blok düzeni için model denklemi:

$$y_{ij} = \mu_{..} + \beta_i + \tau_j + \varepsilon_{ij} \quad , \quad j = 1, 2, \dots, k; i = 1, 2, \dots, n \tag{3.1}$$

eşitliği ile verilir. Burada;

y_{ij}: i.nci blok ve j.nci denemeye ait gözlem (bağımlı değişken) değeri

k: Deneme (Ana faktör düzey) sayısı

n: Blok sayısı

$\mu_{..}$: Genel kitle ortalaması

τ_j: j.nci deneme etkisi, $\left(\tau_j = \mu_{.j} - \mu_{..}\ ; \sum_{j=1}^{k} \tau_j = 0\right)$

β_i: i.nci blok etkisi, $\beta_i = \mu_{i.} - \mu_{..}$ olup, rastgele olduğundan $\beta_i \sim BND\left(0\ ,\ \sigma_\beta^2\right)$ dir.

ε_{ij}: Hata terimi, $\varepsilon_{ij} = y_{ij} - \mu_{..} - \beta_i - \tau_j = y_{ij} - \mu_{i.} - \mu_{.j} + \mu_{..}$ olup, bir rastgele değişken ve $\varepsilon_{ij} \sim BND(0\ ,\sigma_\varepsilon^2)$ dir. Ayrıca ε_{ij}, β_i'den bağımsızdır.

Rastgele blok düzeninde k tane deneme ve n tane blok olduğunda veri düzeni aşağıdaki gibi olacaktır.

BLOK	**DENEME (FAKTÖR-A)**						$\boldsymbol{T_{i.}}$	$\boldsymbol{\overline{Y}_{i.}}$
	A_1	A_2	…	A_j	…	A_k		
$\boldsymbol{B_1}$	y_{11}	y_{12}	…	y_{1j}		y_{1k}	$\boldsymbol{T_{1.}}$	$\boldsymbol{\overline{Y}_{1.}}$
$\boldsymbol{B_2}$	y_{21}	y_{22}	…	y_{2j}		y_{2k}	$\boldsymbol{T_{2.}}$	$\boldsymbol{\overline{Y}_{2.}}$
⋮	⋮	⋮	…	⋮	⋮	⋮	⋮	⋮
$\boldsymbol{B_i}$	y_{i1}	y_{i2}	…	y_{ij}		y_{ik}	$\boldsymbol{T_{i.}}$	$\boldsymbol{\overline{Y}_{i.}}$
⋮	⋮	⋮	…	⋮	…	⋮	⋮	⋮
$\boldsymbol{B_n}$	y_{n1}	y_{n2}	…	y_{nj}	…	y_{nk}	$\boldsymbol{T_{n.}}$	$\boldsymbol{\overline{Y}_{n.}}$
$\boldsymbol{T_{.j}}$	$T_{.1}$	$T_{.2}$	…	$T_{.j}$	…	$T_{.k}$	$T_{..}$	$N = k * n$
$\boldsymbol{\overline{Y}_{.j}}$	$\boldsymbol{\overline{Y}_{.1}}$	$\boldsymbol{\overline{Y}_{.2}}$	**…**	$\boldsymbol{\overline{Y}_{.j}}$	**…**	$\boldsymbol{\overline{Y}_{.k}}$	$\boldsymbol{\overline{Y}_{..}}$	
$\sum_{i=1}^{n} y_{ij}^2$	$\sum_{i=1}^{n} y_{i1}^2$	$\sum_{i=1}^{n} y_{i2}^2$	…	$\sum_{i=1}^{n} y_{ij}^2$	…	$\sum_{i=1}^{n} y_{ik}^2$	$\sum_{j=1}^{k} \sum_{i=1}^{n} y_{ij}^2$	

Eşitlik (3.1) ile verilen modelin parametreleri EKK metodu ile tahmin edilebilir. Bu amaçla mimimize edilecek olan hata kareler toplamı (KT_{Hata}) fonksiyonu;

$$KT_{Hata}=\sum_{j=1}^{k}\sum_{i=1}^{n}\varepsilon_{ij}^2=\sum_{j=1}^{k}\sum_{i=1}^{n}\left(y_{ij}-\mu_{..}-\beta_i-\tau_j\right)^2 \quad (3.2)$$

dir. Bu fonksiyonun parametrelere $\left(\mu_{..}, \beta_i, \tau_j,\ \mu_{i.},\ \mu_{.j}\right)$ göre sırasıyla türevleri alınarak, sıfıra eşitlenmek suretiyle elde edilecek olan denklemlerin çözümünden tahmin ediciler aşağıdaki gibi bulunur.

Parametre	EKK Tahmin Edicisi
$\mu_{..}$	$\overline{Y}_{..}=\frac{1}{N}\sum_{j=1}^{k}\sum_{i=1}^{n}y_{ij}=\frac{T_{..}}{N}$
$\mu_{i.}$	$\overline{Y}_{i.}=\frac{1}{k}\sum_{j=1}^{k}y_{ij}=\frac{T_{i.}}{k}$, $i=1,2,\dots,n$
$\mu_{.j}$	$\overline{Y}_{.j}=\frac{1}{n}\sum_{i=1}^{n}y_{ij}=\frac{T_{.j}}{n}$, $j=1,2,\dots,k$
β_i	$\hat{\beta}_i=\overline{Y}_{i.}-\overline{Y}_{..}$, , $i=1,2,\dots,n$
τ_j	$\hat{\tau}_j=\overline{Y}_{.j}-\overline{Y}_{..}$, $j=1,2,\dots,k$
ε_{ij}	$\hat{\varepsilon}_{ij}=y_{ij}-\overline{Y}_{i.}-\overline{Y}_{.j}+\overline{Y}_{..}$

Eşitlik (3.1) ile verilen modelde denemeler sabit etkili, bloklar rastgele etkili seçildiğinden bu tür rastgele blok düzeni modeline karışım modeli adı verilir. Burada $\varepsilon_{ij}\sim BND(0,\sigma_\varepsilon^2)$ olduğundan, gözlem birimlerinin dağılımı; $y_{ij}=\mu_{..}+\beta_i+\tau_j+\varepsilon_{ij}\sim N\left(\mu_{..}+\beta_i+\tau_j,\sigma_\varepsilon^2\right)$ olacaktır. Rastgele deney düzeninde esas amaç denemelerin (yani faktör düzeylerinin) bağımlı değişken üzerine etkisini incelemektir. Bunun yanı sıra bloklamanın etkisi de incelenebilir. Bu durumda test edilecek hipotezler bu iki etkinin önemliliğini araştırmak üzerine kurulmalıdır.

i) Deneme etkisinin önemliliği için, denemeler sabit etkili olduğundan

$H_0: \tau_1=\tau_2=\cdots=\tau_k=0$ (veya $H_0: \mu_{.1}=\mu_{.2}=\cdots=\mu_{.k}=\mu_{..}$)

$H_1: \exists\tau_j\neq 0$ (veya $H_1: \exists\mu_{.j}$ diğerlerinden farklı (3.3)

ii) Blok etkisinin önemliliği için, bloklar rastgele etkili olduğundan

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

$H_1: \sigma_\beta^2>0$ (3.4)

şeklinde oluşturulur. H_0 hipotezinin test edilmesinde gerekli olan test istatistiklerini türetebilmek için, Eşitlik (3.1) ile verilen modelde parametreler yerine EKK tahmin edicileri kullanılır ve bağımlı değişkene ait toplam değişim (GKT), birbirinden bağımsız olan varyans kaynaklarına (Blok, deneme ve hata) ayrıştırılır.

$$y_{ij}=\mu_{..}+\beta_i+\tau_j+\varepsilon_{ij} \implies y_{ij}-\mu_{..}=\beta_i+\tau_j+\varepsilon_{ij} \implies$$

$$y_{ij}-\overline{Y}_{..}=\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)+\left(\overline{Y}_{.j}-\overline{Y}_{..}\right)+\left(y_{ij}-\overline{Y}_{i.}-\overline{Y}_{.j}+\overline{Y}_{..}\right),\ j=1,2,\dots,k;\ i=1,2,\dots,n$$

Bu eşitliğin sol tarafı “genel sapma”, sağ tarafta birinci terim “Blok etkisi”, ikinci terim “deneme etkisi” ve üçüncü terim “hata” olarak bilinir. Her iki tarafın karesi alındıktan sonra i ve j üzerinden toplam alınırsa;

$$\sum_{i=1}^{n}\sum_{j=1}^{k}\left(y_{ij}-\overline{Y}_{..}\right)^2=\sum_{i=1}^{n}\sum_{j=1}^{k}\left[\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)+\left(\overline{Y}_{.j}-\overline{Y}_{..}\right)+\left(y_{ij}-\overline{Y}_{i.}-\overline{Y}_{.j}+\overline{Y}_{..}\right)\right]^2$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{k}\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)^2+\sum_{i=1}^{n}\sum_{j=1}^{k}\left(\overline{Y}_{.j}-\overline{Y}_{..}\right)^2+\sum_{i=1}^{n}\sum_{j=1}^{k}\left(y_{ij}-\overline{Y}_{i.}-\overline{Y}_{.j}+\overline{Y}_{..}\right)^2$$

$$+\ 2\{3\ tane\ ikili\ çarpım\ terimi\}$$

$$GKT=k\sum_{i=1}^{n}\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)^2+n\sum_{j=1}^{k}\left(\overline{Y}_{.j}-\overline{Y}_{..}\right)^2+\sum_{i=1}^{n}\sum_{j=1}^{k}\left(y_{ij}-\overline{Y}_{i.}-\overline{Y}_{.j}+\overline{Y}_{..}\right)^2$$

$$=KT_{Blok}+KT_{Deneme}+KT_{Hata} \qquad (3.5)$$

elde edilir. Bu kareler toplamları için hesaplama formülleri ise;

$$GKT=\sum_{i=1}^{n}\sum_{j=1}^{k}\left(y_{ij}-\overline{Y}_{..}\right)^2=\sum_{i=1}^{n}\sum_{j=1}^{k}{y_{ij}}^2-\frac{{T_{..}}^2}{N} \qquad (3.6)$$

$$KT_{Deneme}=n\sum_{j=1}^{k}\left(\overline{Y}_{.j}-\overline{Y}_{..}\right)^2=\sum_{j=1}^{k}\frac{T_{.j}^2}{n}-\frac{{T_{..}}^2}{N} \qquad (3.7)$$

$$KT_{Blok}=k\sum_{i=1}^{n}\left(\overline{Y}_{i.}-\overline{Y}_{..}\right)^2=\sum_{i=1}^{n}\frac{T_{i.}^2}{k}-\frac{{T_{..}}^2}{N} \qquad (3.8)$$

$$KT_{Hata}=GKT-KT_{Deneme}-KT_{Blok}$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{k}{y_{ij}}^2-\sum_{i=1}^{n}\frac{T_{i.}^2}{k}-\sum_{j=1}^{k}\frac{T_{.j}^2}{n}+\frac{{T_{..}}^2}{N} \qquad (3.9)$$

şeklinde düzenlenir. Burada KT_{Blok}, $KT_{Deneme}\ ve\ KT_{Hata}$ birbirinden bağımsızdır. Her bir kareler toplamı kendi serbestlik derecesine bölünerek $KO_{Blok}, KO_{Deneme}\ ve\ KO_{Hata}$ istatistikleri hesaplanır.

$$KO_{Deneme}=\frac{KT_{Deneme}}{k-1}\ ;\ KO_{Blok}=\frac{KT_{Blok}}{n-1}\ ;\ KO_{Hata}=\frac{KT_{Hata}}{(n-1)(k-1)} \qquad (3.10)$$

Deneme ve Blok etkilerine ait kareler ortalamaları, KO_{Hata}’ya bölünerek test istatistikleri elde edilir.

Eşitlik (3.3) ile verilen H_0 hipotezini test etmek için test istatistiği;

$$F_{Deneme}=\frac{KO_{Deneme}}{KO_{Hata}}\sim F_{k-1;(n-1)(k-1)} \qquad (3.11)$$

dir. α önem seviyesinde eğer; $F_{Deneme}>F_{k-1;(n-1)(k-1);\alpha}$ ise H_0 hipotezi ret edilir, $F_{Deneme}\le F_{k-1;(n-1)(k-1);\alpha}$ ise H_0 hipotezi ret edilemez.

Eşitlik (3.4) ile verilen H_0 hipotezini test etmek için test istatistiği;

$$F_{Blok} = \frac{KO_{Blok}}{KO_{Hata}} \sim F_{n-1;(n-1)(k-1)} \tag{3.12}$$

dir. α önem seviyesinde eğer; $F_{Blok} > F_{n-1;(n-1)(k-1);\alpha}$ ise H_0 hipotezi ret edilir, $F_{Blok} \leq F_{n-1;(n-1)(k-1);\alpha}$ ise H_0 hipotezi ret edilemez.

<u>SONUÇ:</u> ANOVA TABLOSU

Kaynak	S.D.	KT	KO	Test İstatistiği
Deneme(τ_j)	$k-1$	KT_{Deneme}	KO_{Deneme}	$F_{Deneme} = \frac{KO_{Deneme}}{KO_{Hata}}$
Blok(β_i)	$n-1$	KT_{Blok}	KO_{Blok}	$F_{Blok} = \frac{KO_{Blok}}{KO_{Hata}}$
Hata(ε_{ij})	$(k-1)(n-1)$	KT_{Hata}	KO_{Hata}	……………..
Genel	N-1	KT_{Genel}		…………..

III.2 GENELLEŞTİRİLMİŞ RASTGELE BLOK DÜZENİ (İKİ FAKTÖR VARYANS ANALİZİ)

Rastgele blok düzeninde her blokta her deneme yalnız bir defa kullanılmaktadır. Bu sebeple blok*deneme etkileşimi test etmek mümkün olmamaktadır. Bu etkileşimi test edebilmek için rastgele blok düzeninde şu şartların sağlanması gerekir.

i) Deney düzeninde en az iki düzeyi olan tek bir deneme (faktör) ve en az iki düzeyi olan tek bir gürültü değişkeni vardır. Bu gürültü değişkenine, burada blok değil, grup adı verilecektir. Her bir grupta $n * k$ sayıda homojen gözlem biriminin bulunduğu g tane grup olmalıdır. Ancak; $n > 1$ olmak zorundadır.

ii) $n * k$ sayıdaki gözlem birimlerinin her birine uygulanacak denemeler rastgele belirlenir. Ancak; her bir gruptaki her bir denemeye tam n tane birim atanmak zorundadır. Toplam gözlem sayısı $N = n * k * g$ dir.

Bu şekilde plânlanan deney düzenine **genelleştirilmiş rastgele blok düzeni** veya **iki faktör varyans analizi** adı verilir. Bu deney düzeni için model:

$$y_{ijz} = \mu_{...} + \tau_j + \delta_z + (\tau\delta)_{jz} + \varepsilon_{i(jz)} , \begin{cases} i = 1,2,\dots,n \\ j = 1,2,\dots,k \\ z = 1,2,\dots,g \end{cases} \tag{3.13}$$

şeklinde ifade edilir. Bu modelde;

y_{ijz}: j.nci deneme ve z.nci gruptaki i.nci birimin aldığı değer

$\mu_{...}$: Genel kitle ortalaması

τ_j: j.nci deneme etkisi (özel seçimli olduğundan sabit etkili olup, $\tau_j = \mu_{.j.} - \mu_{...}$; $\sum_{j=1}^{k} \tau_j = 0$)

δ_z: z.nci grubun etkisi (hem sabit hem de rastgele etkili olabilir, özel seçimli olduğunda sabit etkili olup, $\delta_z = \mu_{..z} - \mu_{...}$; $\sum_{z=1}^{g} \delta_z = 0$, rastgele seçimli olduğunda rastgele etkili olup, $\delta_z \sim BND(0\,,\ \sigma_\delta^2)$ dir)

$(\tau\delta)_{jz}$: Etkileşim etkisi [j.nci deneme ile z.nci grubun ortak etkisini ifade eder ve denemelerin sabit, grupların rastgele etkili olduğu bir deneyde etkileşim etkisi de rastgele etkilidir. Bu sebeple $(\tau\delta)_{jz} \sim BND\left(0,\ \frac{k-1}{k}\sigma_{(\tau\delta)}^2\right)$ şeklinde gösterilir. Her ikisi de sabit iken sabit etkilidir. $(\tau\delta)_{jz} = \mu_{.jz} - \mu_{.j.} - \mu_{..z} + \mu_{...}$ şeklinde ifade edilir. Ayrıca bu etkileşim, δ_z'den bağımsız ve $\sum_{j=1}^{k} (\tau\delta)_{jz} = 0$ dır.]

$\varepsilon_{i(jz)}$: Hata terimi olup, bu hem δ_z'den hem de $(\tau\delta)_{jz}$'den bağımsızdır.

Genelleştirilmiş rastgele blok deney düzenine ait veri düzeni ise şu şekildedir:

GRUP	**DENEME (FAKTÖR)**			
	1	**2**	**. . .**	***k***
1	y_{111} y_{211} . . . y_{n11}	y_{121} y_{221} . . . y_{n21}		y_{1k1} y_{2k1} . . . y_{nk1}
2	y_{112} y_{212} . . . y_{n12}	y_{122} y_{222} . . . y_{n22}		y_{1k2} y_{2k2} . . . y_{nk2}
. **.** **.**				
g	y_{11g} y_{21g} . . . y_{n1g}	y_{12g} y_{22g} . . . y_{n2g}		y_{1kg} y_{2kg} . . . y_{nkg}

Bu deney düzeninde test edilecek hipotezler:

a) Deneme etkisinin önemliliği için, denemeler sabit etkili olduğundan

$H_0: \tau_1 = \tau_2 = \cdots = \tau_k = 0$ (veya $H_0: \mu_{.1} = \mu_{.2} = \cdots = \mu_{.k} = \mu_{..}$)

$H_1: \exists \tau_j \neq 0$ (veya $H_1: \exists \mu_{.j}$ diğerlerinden farklı) (3.14)

b) Blok (grup) etkisinin önemliliği için, bloklar rastgele etkili olduğundan

$H_0: \sigma_\delta^2 = 0$ (veya $\delta_1 = \delta_2 = \cdots = \delta_g = 0$)

$H_1: \sigma_\delta^2 > 0$ (veya $\exists z\ için\ \delta_z \neq 0$ (3.15)

c) Etkileşim etkisinin önemliliği için, rastgele etkili olduğundan

$H_0: \sigma_{(\tau\delta)}^2 = 0$ (veya $\tau\delta_{11} = \tau\delta_{12} = \cdots = \tau\delta_{kg} = 0$)

$H_1: \sigma_{(\tau\delta)}^2 > 0$ (veya en az bir (j, z) için $\tau\delta_{jz} \neq 0$) (3.16)

şeklinde oluşturulur. Bu hipotezlerin test edilmesinde gerekli olan test istatistiklerinin elde edilebilmesi için, Genel Kareler Toplamının (GKT), birbirinden bağımsız dört varyans kaynağına ayrıştırılması gerekir. Bu varyans kaynakları "deneme, grup, deneme*grup etkileşimi ve hata" şeklindedir.

$$KT_{Genel} = KT_{Deneme} + KT_{Grup} + KT_{Deneme*Grup} + KT_{Hata} \quad (3.17)$$

denklemi gereğince, genel kareler toplamı birbirinden bağımsız dört parçaya ayrılacaktır.

$$KT_{Genel} = \sum_{j=1}^{k}\sum_{z=1}^{g}\sum_{i=1}^{n}\left(y_{ijz} - \overline{Y}_{...}\right)^2 = \sum_{j=1}^{k}\sum_{z=1}^{g}\sum_{i=1}^{n} {y_{ijz}}^2 - \frac{T_{...}^2}{N} \quad (3.18)$$

$\overline{Y}_{...} = \frac{1}{N}\sum_{j=1}^{k}\sum_{z=1}^{g}\sum_{i=1}^{n} y_{ijz} = \frac{T_{...}}{N} \Longrightarrow \mu_{...}$ parametresinin EKK tahmin edicisi

$$KT_{Deneme} = \sum_{j=1}^{k} gn\left(\overline{Y}_{.j.} - \overline{Y}_{...}\right)^2 = \sum_{j=1}^{k}\frac{T_{.j.}^2}{gn} - \frac{T_{...}^2}{N} \quad (3.19)$$

$\overline{Y}_{.j.} = \frac{1}{gn}\sum_{z=1}^{g}\sum_{i=1}^{n} y_{ijz} = \frac{T_{.j.}}{gn}$ $, j = 1, 2, \dots, k \Longrightarrow \mu_{.j.}$ parametresinin EKK tahmin edicisi

$$KT_{Grup} = \sum_{z=1}^{g} kn\left(\overline{Y}_{..z} - \overline{Y}_{...}\right)^2 = \sum_{z=1}^{g}\frac{T_{..z}^2}{kn} - \frac{T_{...}^2}{N} \quad (3.20)$$

$\overline{Y}_{..z} = \frac{1}{kn}\sum_{j=1}^{k}\sum_{i=1}^{n} y_{ijz} = \frac{T_{..z}}{kn}$ $, z = 1, 2, \dots, g \Longrightarrow \mu_{..z}$ parametresinin EKK tahmin edicisi

$$KT_{Deneme*Grup} = \sum_{j=1}^{k}\sum_{z=1}^{g} n\left(\overline{Y}_{.jz} - \overline{Y}_{.j.} - \overline{Y}_{..z} + \overline{Y}_{...}\right)^2$$

$$= \sum_{j=1}^{k}\sum_{z=1}^{g}\frac{T_{.jz}^2}{n} - \sum_{j=1}^{k}\frac{T_{.j.}^2}{gn} - \sum_{z=1}^{g}\frac{T_{..z}^2}{kn} + \frac{T_{...}^2}{N}$$

$$= \sum_{j=1}^{k}\sum_{z=1}^{g}\frac{T_{.jz}^2}{n} - KT_{Deneme} - KT_{Grup} - \frac{T_{...}^2}{N} \quad (3.21)$$

$\overline{Y}_{.jz} = \frac{1}{n}\sum_{i=1}^{n} y_{ijz} = \frac{T_{.jz}}{n}$, $j = 1, 2, \dots, k;\ z = 1, 2, \dots, g \Longrightarrow \mu_{.jz}$ parametresinin EKK tahmin edicisi

Böylece Eşitlik (3.17) den;

$$KT_{Hata} = KT_{Genel} - KT_{Deneme} - KT_{Grup} - KT_{Deneme*Grup}$$

$$= \sum_{j=1}^{k}\sum_{z=1}^{g}\sum_{i=1}^{n}\left(y_{ijz} - \overline{Y}_{.\ j\ z}\right)^2 = \sum_{j=1}^{k}\sum_{z=1}^{g}\sum_{i=1}^{n} {y_{ijz}}^2 - \sum_{j=1}^{k}\sum_{z=1}^{g}\frac{T^2_{.\ j\ z}}{n} \tag{3.22}$$

elde edilir. Diğer işlemler ve test istatistikleri Anova Tablosunda özetlenebilir.

ANOVA TABLOSU

Kaynak	**S.D.**	**KT**	**KO**	**Test İstatistiği**
Deneme	$k-1$	KT_{Deneme}	$KO_{Deneme} = \frac{KT_{Deneme}}{k-1}$	F_{Deneme}
Grup	$g-1$	KT_{Grup}	$KO_{Grup} = \frac{KT_{Grup}}{g-1}$	F_{Grup}
Deneme*Grup	$(k-1)*(g-1)$	KT_{Deneme}	$KO_{Deneme*Grup} = \frac{KT_{Deneme*Grup}}{(k-1)(g-1)}$	$F_{Deneme*Grup}$
Hata	$kg(n-1)$	KT_{Hata}	$KO_{Hata} = \frac{KT_{Hata}}{kg(n-1)}$	…………………..
Genel	$N-1$	KT_{Genel}	…………………..	………………….

<u>**Karar**</u> α önem seviyesinde eğer;

$F_{Deneme} > F_{k-1;kg(n-1);\alpha} \Longrightarrow$ (3.14) deki H_0 ret edilir

$F_{Deneme} \leq F_{k-1;kg(n-1);\alpha} \Longrightarrow$ (3.14) deki H_0 ret edilemez.

$F_{Grup} > F_{g-1;kg(n-1);\alpha} \Longrightarrow$ (3.15) deki H_0 ret edilir

$F_{Grup} \leq F_{g-1;kg(n-1);\alpha} \Longrightarrow$ (3.15) deki H_0 ret edilemez

$F_{Deneme*Grup} > F_{(k-1)(g-1);kg(n-1);\alpha} \Longrightarrow$ (3.16) daki H_0 ret edilir

$F_{Deneme*Grup} \leq F_{(k-1)(g-1);kg(n-1);\alpha} \Longrightarrow$ (3.16) daki H_0 ret edilemez.

<u>**ÖRNEK:3.1**</u> Dört farklı cerrahi yöntemin kullanıldığı, 6 farklı ağırlık arasından rastgele seçilen üç farklı ağırlık grubundaki (m_1: $zayıf, m_2$: $orta\ ve\ m_3$: $kilolu$) hastaların iyileşme sürelerine (gün) ilişkin veriler aşağıdaki gibi gözlenmiştir. Bu çalışma için;

a) En uygun olan istatistiksel analizi ve model denklemini belirleyiniz?

b) Model parametreleri için EKK tahmin edicilerini bulunuz?

c) ANOVA tablosunu düzenleyiniz?

d) Cerrahi yöntemin iyileşme süresi üzerine etkisinin önemli olup olmadığını belirleyiniz?

e) Hasta ağırlığının iyileşme süresi üzerine etkisinin önemli olup olmadığını belirleyiniz?

f) Eğer gerekiyorsa cerrahi yöntem ve hasta ağırlığı için çoklu karşılaştırma ile farklılık gösteren ikili kombinasyonları belirleyiniz?

		Cerrahi Yöntem				
		1-nci Yöntem	**2-nci Yöntem**	**3-nci Yöntem**	**4-nci Yöntem**	$T_{i,}$
Hastaların Ağırlığı	m_1	13	12	9	10	44
	m_2	12	13	12	11	48
	m_3	15	14	10	11	50
$T_{.j}$		40	39	31	32	$T_{..} = 142$
$\sum_{i=1}^{n} y_{ij}^2$		**538**	**509**	**325**	**342**	$\sum_{j=1}^{k}\sum_{i=1}^{n} y_{ij}^2 = 1714$

<u>Çözüm</u> **a)** Bağımlı değişken (Y): İyileşme süresi (gün)… Nicel, sürekli ve ölçme düzeyi oranlama

Faktör (Deneme): Cerrahi Yöntemi… Nitel ve ölçme düzeyi sınıflama

Denemeler: Birinci cerrahi yöntemi, İkinci cerrahi yöntemi, Üçüncü cerrahi yöntemi, Dördüncü cerrahi yöntemi olup ($k = 4$)

Blok: Hastanın ağırlığı… Nicel, sürekli ve ölçme düzeyi sıralama

Bloklar: $m_1: zayıf, m_2: orta\ ve\ m_3: kilolu$ olup, ($n = 3$)

Uygun İstatistiksel analiz: Denemeler özel seçimli ve bloklar rastgele seçimli olduğundan Rastgele blok düzenidir. Uygun model ise karışım modeli olup Rastgele Blok tasarımıdır. Model denklemi;
$y_{ij} = \mu_{..} + \beta_i + \tau_j + \varepsilon_{ij}$, $j = 1, 2, \ldots, k; i = 1, 2, \ldots, n$ şeklindedir. Burada $\tau_j = \mu_{.j} - \mu_{..}$ ve $\beta_i = \mu_{i.} - \mu_{..}$ $k = 4, n = 3$'dür.

b)

Parametre	EKK Tahmin edicisi	**Parametre**	EKK Tahmin edicisi
$\mu_{..}$	$\overline{Y}_{..} = 142/12=11{,}83$	τ_1	$\hat{\tau}_1 = \overline{Y}_{.1} - \overline{Y}_{..}$=13,33-11,83=1,5
$\mu_{.1}$	$\overline{Y}_{.1} = 40/3=13{,}33$	τ_2	$\hat{\tau}_2 = \overline{Y}_{.2} - \overline{Y}_{..} = 13 - 11{,}83 = 1{,}16$
$\mu_{.2}$	$\overline{Y}_{.2}$=39/3=13	τ_3	$\hat{\tau}_3$=$\overline{Y}_{.3} - \overline{Y}_{..}$=10,33-11,83= -1,5
$\mu_{.3}$	$\overline{Y}_{.3} = 31/3=10{,}33$	τ_4	$\hat{\tau}_4 = \overline{Y}_{.4} - \overline{Y}_{..}$=10,67-11,83= -1,16
$\mu_{.4}$	$\overline{Y}_{.4} = 32/3=10{,}67$	β_1	$\hat{\beta}_1 = \overline{Y}_{1.} - \overline{Y}_{..} = 11 - 11{,}83 = -0{,}83$
$\mu_{1.}$	$\overline{Y}_{1.} = 44/4=11$	β_2	$\hat{\beta}_2$=$\overline{Y}_{2.} - \overline{Y}_{..} = 12 - 11{,}83 = 0{,}17$
$\mu_{2.}$	$\overline{Y}_{2.} = 48/4=12$	β_3	$\hat{\beta}_3 = \overline{Y}_{3.} - \overline{Y}_{..} = 12{,}5 - 11{,}83 = 0{,}67$
$\mu_{3.}$	$\overline{Y}_{3.} = 50/4=12{,}5$		

c) ANOVA TABLOSU

Kaynak	S.D.	KT	KO	Test İstatistiği
Deneme(τ_j)	$k-1=3$	21,667	7,222	$F_{Deneme}=\frac{KO_{Deneme}}{KO_{Hata}}$ =5,91
Blok(β_i)	$n-1$=2	4,667	2,334	$F_{Blok}=\frac{KO_{Blok}}{KO_{Hata}}$ =1,91
Hata(ε_{ij})	$(k-1)(n-1)=6$	7,333	1,222	……………..
Genel	N-1=11	33,667		…………..

$KT_{Genel}=\sum_{j=1}^{k}\sum_{i=1}^{n}y_{ij}^2-\frac{T_{..}^2}{N}=1714-\frac{(142)^2}{12}=33,667$

$KT_{Deneme}=\sum_{j=1}^{k}\frac{T_{.j}^2}{n}-\frac{T_{..}^{\;2}}{N}=\frac{(40)^2}{3}+\frac{(39)^2}{3}+\frac{(31)^2}{3}+\frac{(32)^2}{3}-\frac{(142)^2}{12}=21,667$

$KT_{blok}=\sum_{i=1}^{n}\frac{T_{i.}^2}{k}-\frac{T_{..}^{\;2}}{N}=\frac{(44)^2}{4}+\frac{(48)^2}{4}+\frac{(50)^2}{4}-\frac{(142)^2}{12}=4,667$

$KT_{Hata}=KT_{Genel}-KT_{Deneme}-KT_{blok}=33,667-21,667-4,667=7,333$

d) Cerrahi yöntemin iyileşme süresi üzerine etkisinin önemliliği için test edilecek hipotezler:

$H_0: \mu_{.1}=\mu_{.2}=\mu_{.3}=\mu_{.4}=\mu_{..}$

$H_1: \exists\ \mu_{.j}$ diğerlerinden farklı

şeklinde oluşturulur. Test istatistiği $F_{Deneme}=\frac{KO_{Deneme}}{KO_{Hata}}\sim F_{k-1;\,(k-1)(n-1)}$ dir. H_0 doğru iken test istatistiğinin alabileceği değer, Anova tablosuna göre $F_{Deneme}=\frac{7,222}{1,222}=5,91$ olarak bulunur.

Karar: $\alpha=0,05$ önem seviyesinde H_1hipotezine göre kritik değer $F_t=F_{k-1:(n-1)(k-1);\alpha}$ olmak üzere karar kuralı $F_{Deneme}>F_t$ ise H_0 ret edilir, aksi takdirde ret edilemez.

$F_t=F_{k-1:(n-1)(k-1);\alpha}=F_{3;6;0,05}=4,76$ ve 5,91>4,76 olduğundan H_0 ret edilir. Buna göre cerrahi yöntemlere göre hastaların ortalama iyileşme süreleri istatistiksel olarak 0,05 önem seviyesinde anlamlı bir farklılık göstermektedir. Diğer bir ifadeyle 0,05 önem seviyesinde cerrahi yöntemlerin 3 farklı ağırlık grubundaki hastaların iyileşme sürelerine etkisi önemlidir.

e) Hasta ağırlığının iyileşme süresi üzerine etkisinin önemliliği için test edilecek hipotezler, bloklar rastgele olduğundan;

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

$H_1: \sigma_\beta^2>0$ şeklinde kurulur. Test istatistiği $F_{Blok}=\frac{KO_{Blok}}{KO_{Hata}}\sim F_{n-1;\,(k-1)(n-1)}$ dir. H_0 doğru iken test istatistiğinin alabileceği değer, Anova tablosuna göre $F_{Blok}=\frac{2,334}{1,222}=1,91$ olarak bulunur.

Karar: $\alpha=0,05$ önem seviyesinde H_1hipotezine göre kritik değer $F_t=F_{n-1:(n-1)(k-1);\alpha}$ olmak üzere karar kuralı $F_{Blok}>F_t$ ise H_0 ret edilir, aksi takdirde ret edilemez.

$F_t = F_{n-1:(n-1)(k-1);\alpha} = F_{2;6;0,05} = 5{,}14$ ve 1,91<5,14 olduğundan H_0 ret edilemez. Buna göre 0,05 önem seviyesinde bloklar (yani üç farklı ağırlık gurubu) arasındaki farkın hastaların iyileşme süresi üzerinde etkisinin önemli olmadığı söylenir.

f) Bloklar arasında hastaların iyileşme süreleri bakımından farklılık olmadığı için bloklarla ilgili ikili karşılaştırma yapılmasına gerek yoktur. Cerrahi yöntemler (Denemeler) arasında anlamlı farklılık bulunduğundan farklılık gösteren denemeleri belirlemek için çoklu karşılaştırmalara başvurulmalıdır.

i) LSD çoklu karşılaştırma testini uygulayalım. Mümkün olan ikili karşılaştırmaların sayısı $\binom{k}{2} = \binom{4}{2} = 6$ dır.

a) 1.Yöntem-2.Yöntem karşılaştırması:

$H_0: \mu_{.\,1} = \mu_{.\,2}$

$H_1: \mu_{.\,1} \neq \mu_{.\,2}$

Karşılaştırma kriteri $\left|\overline{Y}_{.\,1} - \overline{Y}_{.\,2}\right|$ >LSD ise H_0 ret edilir. $n_1 = n_2 = n_3 = n_4 = n = 3$ olduğundan

$\overline{Y}_{1.} = 13{,}33$ ve $\overline{Y}_{2.} = 13$ olup, $\left|\overline{Y}_{.\,1} - \overline{Y}_{.\,2}\right| = 0{,}33$ ve

$$LSD = t_{\alpha/2, s.d._{hata}} * \sqrt{\frac{2KO_{Hata}}{n}} = t_{0,025;6} * \sqrt{\frac{2*1{,}222}{3}} = (2{,}447) * (0{,}903) = 2{,}21$$

bulunur. Buna göre; $\left|\overline{Y}_{.\,1} - \overline{Y}_{.\,2}\right| = 0{,}33 < 2{,}213 = LSD$ olduğundan H_0 ret edilemez. Yani; 1.Yöntem ile 2.Yöntem arasında %95 güvenle anlamlı bir farklılık yoktur.

b) 1.Yöntem-3.Yöntem karşılaştırması:

$H_0: \mu_{.\,1} = \mu_{.\,3}$

$H_1: \mu_{.\,1} \neq \mu_{.\,3}$

Karşılaştırma kriteri $\left|\overline{Y}_{.\,1} - \overline{Y}_{.\,3}\right| > LSD$ ise H_0 ret edilir. $LSD = 2{,}21$, $\left|\overline{Y}_{.\,1} - \overline{Y}_{.\,3}\right| = |13{,}33 - 10{,}33| = 3 > 2{,}21$ olduğundan H_0 ret edilir. Yani; 1.Yöntem ile 3.Yöntem arasında %95 güvenle anlamlı bir farklılık vardır.

c) 1.Yöntem-4.Yöntem karşılaştırması:

$H_0: \mu_{.\,1} = \mu_{.\,4}$

$H_1: \mu_{.\,1} \neq \mu_{.\,4}$

Karşılaştırma kriteri $\left|\overline{Y}_{.\,1} - \overline{Y}_{.\,4}\right| > LSD$ ise H_0 ret edilir. $LSD = 2{,}21$, $\left|\overline{Y}_{.\,1} - \overline{Y}_{.\,4}\right| = |13{,}33 - 10{,}67| = 2{,}66 > 2{,}21$ olduğundan H_0 ret edilir. Yani; 1.Yöntem ile 4.Yöntem arasında %95 güvenle anlamlı bir farklılık yoktur.

d) 2.Yöntem-3.Yöntem karşılaştırması:

H_0: $\mu_{.\,2} = \mu_{.\,3}$

H_1: $\mu_{.\,2} \neq \mu_{.\,3}$

Karşılaştırma kriteri $\left|\overline{Y}_{.\,2} - \overline{Y}_{.\,3}\right| > LSD$ ise H_0 ret edilir. $LSD = 2{,}21$, $\left|\overline{Y}_{.\,2} - \overline{Y}_{.\,3}\right| = |13 - 10{,}33| = 2{,}67 > 2{,}21$ olduğundan H_0 ret edilir. Yani; 2.Yöntem ile 3.Yöntem arasında %95 güvenle anlamlı bir farklılık vardır.

e) 2.Yöntem-4.Yöntem karşılaştırması:

H_0: $\mu_{.\,2} = \mu_{.\,4}$

H_1: $\mu_{.\,2} \neq \mu_{.\,4}$

Karşılaştırma kriteri $\left|\overline{Y}_{.\,2} - \overline{Y}_{.\,4}\right| > LSD$ ise H_0 ret edilir. $LSD = 2{,}21$, $\left|\overline{Y}_{.\,2} - \overline{Y}_{.\,4}\right| = |13 - 10{,}67| = 2{,}33 > 2{,}21$ olduğundan H_0 ret edilir. Yani; 2.Yöntem ile 4.Yöntem arasında %95 güvenle anlamlı bir farklılık vardır.

f) 3.Yöntem-4.Yöntem karşılaştırması:

H_0: $\mu_{.\,3} = \mu_{.\,4}$

H_1: $\mu_{.\,3} \neq \mu_{.\,4}$

Karşılaştırma kriteri $\left|\overline{Y}_{.\,3} - \overline{Y}_{.\,4}\right| > LSD$ ise H_0 ret edilir. $LSD = 2{,}21$, $\left|\overline{Y}_{.\,3} - \overline{Y}_{.\,4}\right| = |10{,}33 - 10{,}67| = 0{,}34 < 2{,}21$ olduğundan H_0 ret edilemez. Yani; 3.Yöntem ile 4.Yöntem arasında %95 güvenle anlamlı bir farklılık yoktur.

ii) Newman-Keuls çoklu karşılaştırma testini uygulayalım. Mümkün olan ikili karşılaştırmaların sayısı $\binom{k}{2} = \binom{4}{2} = 6$ dır. Hipotezler;

H_0: $\mu_{.\,j} = \mu_{.\,t}$

H_1: $\mu_{.\,j} \neq \mu_{.\,t}$, $j \neq t = 1,2,3,4$

Grup No:	1	2	3	4
Ortalama$(\overline{Y}_{j.})$:	13,33	13	10,33	10,67

Grup No:	1	2	4	3
Sıralı Ortalama$(\overline{Y}_{j.})$:	13,33	13	10,67	10,33

Karşılaştırma kriteri: $\overline{Y}_{j.} > \overline{Y}_{t.}$ olmak üzere $\overline{Y}_{j.} - \overline{Y}_{t.} > W_r$, r=2, 3, 4 (k=4 olduğundan) ise H_0 ret edilir.

$W_r = q_{r;s.d._{hata};\alpha}\sqrt{\frac{KO_{Hata}}{n}}$, $r = 2,3,4; n = 3$; $KO_{Hata} = 1,222$; $s.d._{hata} = 6$

r	$q_{r;6;0,05}$	W_r
2	3,46	2,208
3	4,34	2,770
4	4,90	3,127

$\overline{Y}_{1.} - \overline{Y}_{2.} = 13,33 - 13 = 0,33 < 2,208 = W_2$ olduğundan $H_0: \mu_{.1} = \mu_{.2}$ ret edilemez

$\overline{Y}_{1.} - \overline{Y}_{4.} = 13,33 - 10,67 = 2,66 < 2,77 = W_3$ olduğundan $H_0: \mu_{.1} = \mu_{.4}$ ret edilemez

$\overline{Y}_{1.} - \overline{Y}_{3.} = 13,33 - 10,33 = 3,00 < 3,127 = W_4$ olduğundan $H_0: \mu_{.1} = \mu_{.3}$ ret edilemez

$\overline{Y}_{2.} - \overline{Y}_{4.} = 13 - 10,67 = 2,33 > 2,208 = W_2$ olduğundan $H_0: \mu_{.2} = \mu_{.4}$ ret edilir

$\overline{Y}_{2.} - \overline{Y}_{3.} = 13 - 10,33 = 2,67 < 2,77 = W_3$ olduğundan $H_0: \mu_{.2} = \mu_{.3}$ ret edilemez

$\overline{Y}_{4.} - \overline{Y}_{3.} = 10,67 - 10,33 = 0,34 < 2,208 = W_2$ olduğundan $H_0: \mu_{.4} = \mu_{.3}$ ret edilemez

Newman-Keuls testine göre sadece 2-nci ve 4-ncü yöntemler arasında anlamlı bir farklılık var. Diğer tüm ikili karşılaştırmalar için ortalamalar arasında anlamlı bir fark bulunamamıştır.