İlişkili k örneklem-Friedman F Testi

Nihat Tak

2023-05-27

İlişkili k Örneklem Friedman F Testi

- Gruplar arası farklılığı belirlerken; aynı gruptaki gözlemler arası değişkenlik, gruplar arasında olabilecek farkı maskeleyebilir. Bu durumda ilgilenilen değişkendeki grup farklarını belirlemek için, gözlemler homojen alt gruplara bölünebilir. Bu alt gruplara BLOK denir ve karşılaşmalar bloklar arasında yapılır.

- Bu durumda grup içi değişkenliği azalarak, gruplar arası değişkenliği görülmesi kolaylaştırırız.

- Bu test parametrik tekrarlı varyans analizinin parametrik olmayan karşılığıdır.

- Bu testin amacı denemeler arasında farklılığın olup olmadığını ortaya koymaktır.

- Blok ve grup sayısı 5' ten büyük olduğunda iyi sonuç veren testtir.

Parametrik varyans analizinde

- Hatalar birbirinden bağımsız olarak normal dağılmış kitlelerden alınır.
- Populasyonların hepsi aynı varyansa sahip olmalı
- Etkiler toplanabilir olmalıdır.
- Veriler en azından aralıklı ölçekle ölçülmelidir.

Yukarıdaki varsayımlardan biri veya birkaçı sağlanmaz ise parametrik varyans analizi yerine Friedman F testi kullanılır.

Varsayımları

- Gözlemler birbirinden bağımsız m tane bağlantılı gruptan oluşmalıdır.
- Her blokta elde edilen ölçümler en az sıralayıcı ölçekte ölçülmelidir.(Testin gerekleşmesi için verilerin sıralanabilir. olması gerekmektedir.)
- İlgilenilen değişken sürekli olmalıdır.
- Bloklar ve denemeler arası etkileşim olmamalıdır.
- Gruplar aynı değişkenliktedir.

Hipotez testi aşamaları

1. Adım Hipotezler kurulur.

$H_0: M_1 = M_2 \ldots = M_k$ (M=Medyan)

$H_0: \tau_1 = \tau_2 = \cdots = \tau_k$ (verilerin sıralanışı benzer olup, deneme etkileri farklı değildir.)

H_a: En az bir M_j farklıdır $j = 1, \ldots k$ H_a: En az bir τ_i farklıdır (verilerin sıralanışı benzer olmayıp, deneme etkileri farklıdır.)

2. Adım Test istatistiği hesaplanır.

- Öncelikle m tane blok ve her blokta k tane deneme yapılırsa elde edeceğimiz tablo aşağıdaki gibidir.

	Denemeler			
Blok	1	2	......	k
1	x_{11}	x_{12}	....	x_{1k}
2	x_{21}			
....				
m	x_{m1}			x_{mk}

- Her bloktaki gözlemler için sıra sayıları(rank) yazılır ve her blok için sıra sayılarının(rankların) toplamı alınır.

	1	2	......	k	
1	R_{11}	R_{12}	....	R_{1k}	k(k+1)/2
2	R_{21}				k(k+1)/2
....					...
m	R_{m1}			R_{mk}	k(k+1)/2
R_j	**R_1**	**R_2**	**......**	**R_k**	**mk(k+1)/2**

R_j: j. denemenin bloklar toplamı

$$FR = \frac{12}{km(k+1)} \sum_{j=1}^{k} R_j^2 - 3m(k+1)$$

Eğer her blokta gözlemlerin sıra değerleri benzer ise hesapladığımız FR istatistiğinin hesabı değişir. Bu durumda aşağıdaki FR istatistiğini kullanırız.

$$FR = \frac{\frac{12}{km(k+1)} \sum_{j=1}^{k} R_j^2 - 3m(k+1)}{1 - \frac{\sum_{i=1}^{n} T_i}{mk(k^2-1)}}$$

$$\sum T_i = \sum t_i^3 - \sum t_i$$

t_i değeri i. blokta birbirinin aynı sıra numarasını alan gözlem sayısı

ti değeri i. blokta birbirinin aynı sıra numarasını alan gözlem sayısı

3. Adım Kritik tablo değeri bulunur.

$\chi^2_{k-1,\alpha}$ tablo değeri bulunur.

4. Adım Karar verilir.

Eğer $FR \geq \chi^2_{k-1,\alpha}$ ise H_0 reddedilir.

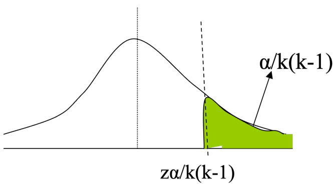
Eğer H_0 reddedilirse denemeler arası fark vardır veya deneme etkileri farklıdır denilir.

Farklı Grupların Saptanması

- $\left|R_{.j} - R_{.j'}\right|$ farklarının hangisi istatistiksel bakımından önemli olduğunun normal yaklaşımı, mevcut grupların sıra numarası toplamları arasındaki farklar ikişer ikişer ele alınarak karşılaştırma yapılmaktadır. Bu fark d_c den büyük ise tesadüfi değil de gerçek bir fark olduğu söylenebilir.

$$d_c = z\sqrt{\frac{mk(k+1)}{6}}$$

- z değeri: z tablosundan $z_{\alpha/k(k-1)}$ e karşılık gelen değerdir.



Örnek Bir bisküvi fabrikasında aynı hammaddeye A:Portakal, B:Kakao, C:Krema, D:Fındık, E:Yulaf ekleyerek 5 ayrı lezzet bisküviyi tüketiciye sunmaya hazırlanıyor. Üretilen prototip bisküviler 8 tadıcı eleman tarafından beğeni testine alınıyor. Her elemanın 5 çeşidin her birine verdikleri 10 üzerinden puanlar şöyle dağılıyor:

	KATKILAR				
DENEMECİ	**A**	**B**	**C**	**D**	**E**
1	8	5	6	7	4
2	10	8	8	4	2
3	9	4	5	6	1
4	9	8	7	5	4
5	10	5	9	8	5
6	7	9	7	5	3
7	8	6	10	4	3
8	8	8	9	7	2

Bu verilere göre “katkılar lezzette farklı etkiler yaratmıştır” denilebilinir mi? Lezzet bakımından A, B, C, D, E katkıları birbirinden farklı mıdır?

1. Adım Hipotezler kurulur.

$H_0: M_1 = M_2 = M_3 = M_4 = M_5$ (Katkı grupları arasında lezzet puanları bakımından önemli bir fark yoktur.)

H_a: En az bir M_j farklıdır.

2. Adım Test istatistiği hesaplanır.

		KATKILAR									
		A		B		C		D		E	
DENEMECİ	1	8	5	5	2	6	3	7	4	4	1
	2	10	5	8	3,5	8	3,5	4	2	2	1
	3	9	5	4	2	5	3	6	4	1	1
	4	9	5	8	4	7	3	5	2	4	1
	5	10	5	5	1,5	9	4	8	3	5	1,5
	6	7	3,5	9	5	7	3,5	5	2	3	1
	7	8	4	6	3	10	5	4	2	3	1
	8	8	3,5	8	3,5	9	5	7	2	2	1
			R_1=36		R_2=24,5		R_3=31		R_4=20		R_5=8,5

$$T_2 = \sum_{i=1}^{1} t_i^3 - \sum_{i=1}^{1} t_i = 2^3 - 2 = 6$$

2. blokta aynı sıra numarasına sahip gözlem değerimiz 2 tanedir.(3,5 sıra numarası alan 2 tane gözlemimiz var)

$$T_5 = 2^3 - 2 = 6$$

5. blokta aynı sıra numarasına sahip gözlem değerimiz 2 tanedir(1,5 sıra numarası alan 2 tane gözlemimiz var)

$$T_6 = 2^3 - 2 = 6$$

$$T_8 = 2^3 - 2 = 6$$

$$\sum_{i=1}^{8} Ti = 6 + 6 + 6 + 6 = 24$$

$$\mathbf{FR} = \frac{\frac{12}{mk(k+1)}\left[\sum_{j=1}^{k} R_j^2 - 3m(k+1)\right]}{1 - \frac{\sum_{i=1}^{n} Ti}{mk(k^2-1)}} = \frac{\frac{12}{8\times5\times(5+1)}\left[3329{,}5 - 3\times8\times(5+1)\right]}{1 - \frac{24}{8\times5\times(5^2-1)}} = 23{,}08$$

3. Adım Kritik tablo değeri bulunur.

$$\chi^2_{k-1,\alpha} = \chi^2_{5-1,0.05} = 9.488$$

4. Adım Karar verilir.

$Fr = 23.08 > 9.488 = \chi^2_{4,0.05}$ olduğundan H_0 reddedilir. Yani 5 ayrı katkı maddesinin sağladığı tatlar arasında tesadüf ile açıklanamayacak düzeyde farklılıklar mevcuttur.

```
x<-
c(8,10,9,9,10,7,8,8,5,8,4,8,5,9,6,8,6,8,5,7,9,7,10,9,7,4,6,5,8,5,4,7,4,2,1,4,
```

```
5,3,3,2)
blok<-rep(1:8,5)
km<-c(rep("A",8),rep("B",8),rep("C",8),rep("D",8),rep("E",8))
(datan<-data.frame(blok,km,x))
```

```
##    blok km  x
## 1     1  A  8
## 2     2  A 10
## 3     3  A  9
## 4     4  A  9
## 5     5  A 10
## 6     6  A  7
## 7     7  A  8
## 8     8  A  8
## 9     1  B  5
## 10    2  B  8
## 11    3  B  4
## 12    4  B  8
## 13    5  B  5
## 14    6  B  9
## 15    7  B  6
## 16    8  B  8
## 17    1  C  6
## 18    2  C  8
## 19    3  C  5
## 20    4  C  7
## 21    5  C  9
## 22    6  C  7
## 23    7  C 10
## 24    8  C  9
## 25    1  D  7
## 26    2  D  4
## 27    3  D  6
## 28    4  D  5
## 29    5  D  8
## 30    6  D  5
## 31    7  D  4
## 32    8  D  7
## 33    1  E  4
## 34    2  E  2
## 35    3  E  1
## 36    4  E  4
## 37    5  E  5
## 38    6  E  3
## 39    7  E  3
## 40    8  E  2
```

```
friedman.test(y=datan$x,groups = datan$km,blocks = datan$blok)
```

```
##
##  Friedman rank sum test
```

```
##
## data:  datan$x, datan$km and datan$blok
## Friedman chi-squared = 22.026, df = 4, p-value = 0.0001981

if(!require(PMCMRplus)) {install.packages("PMCMRplus"); require(PMCMRplus)}

## Loading required package: PMCMRplus

friedmanTest(data=datan,y=datan$x,groups = datan$km,blocks = datan$blok)

##
##  Friedman rank sum test
##
## data:  y, groups and blocks
## Friedman chi-squared = 22.026, df = 4, p-value = 0.0001981

frdAllPairsConoverTest(x,km,blok)

##
##  Pairwise comparisons using Conover's all-pairs test for a two-way
balanced complete block design

## data: y, groups and blocks

##   A       B       C       D
## B 0.37766 -       -       -
## C 0.88317 0.91229 -       -
## D 0.13278 0.98249 0.62564 -
## E 0.00018 0.09183 0.00718 0.29153

##
## P value adjustment method: single-step
```

Örnek Üç ilacın insanlardaki reaksiyon zamanı üzerine etkisi araştırılmaktadır. Bunun için yapılan çalışmanın bulguları aşağıdaki tabloda verilmiştir. Bu durumda bu üç ilacın reaksiyon etkisinin farklı olduğu söylene bilinir mi?

		İLAÇLAR		
		A	**B**	**C**
DENEMECİ	**1**	10	10	15
	2	10	15	20
	3	11	15	12
	4	8	12	10
	5	7	12	9
	6	15	10	15
	7	14	12	18
	8	10	14	17
	9	9	9	12
	10	10	14	16

1.Adım Hipotezler kurulur.

$H_0: M_1 = M_2 = M_3$ (Üç ilacın reaksiyon zamanları arasında fark yoktur.)

H_a: En az bir M_j farklıdır.(Üç ilacın reaksiyon zamanları arasında fark vardır.)

2. Adım Test istatistiği hesaplanır.

		İLAÇLAR					
		A		**B**		**C**	
DENEMECİ	**1**	10	**(1,5)**	10	**(1,5)**	15	**(3)**
	2	10	**(1)**	15	**(2)**	20	**(3)**
	3	11	**(1)**	15	**(2)**	16	**(3)**
	4	8	**(1)**	12	**(3)**	10	**(2)**
	5	7	**(1)**	12	**(2)**	13	**(3)**
	6	15	**(2)**	15	**(2)**	15	**(2)**
	7	14	**(2)**	12	**(1)**	18	**(3)**
	8	10	**(1)**	14	**(2)**	17	**(3)**
	9	9	**(1,5)**	9	**(1,5)**	12	**(3)**
	10	10	**(1)**	14	**(2)**	16	**(3)**
			R_1=13		**R_2=19**		**R_3=28**

$$\sum_{i=1}^{1} t_i^3 - \sum_{i=1}^{1} t_i = 2^2 - 2 = 6$$

$$T_2 = 3^3 - 3 = 24$$

$$T_3 = 2^3 - 2 = 6$$

$$\sum_{i=1}^{3} T_i = 6 + 24 + 6 = 36$$

$$FR = \frac{\frac{12}{km(k+1)} \sum_{j=1}^{k} R_j^2 - 3m(k+1)}{1 - \frac{\sum_{i=1}^{n} T_i}{mk(k^2-1)}}$$

$$FR = \frac{\frac{12}{10 * 3(3+1)} 1314 - 3 * 10(4)}{1 - \frac{36}{10 * 3(3^2-1)}} = \frac{11.4}{0.86} = 13.25$$

3. Adım Kritik tablo değeri bulunur.

$$\chi^2_{k-1,\alpha} = \chi^2_{3-1,0.1} = 4.605$$

4. Adım Karar verilir.

$Fr = 13.25 > 4.605 = \chi^2_{2,0.1}$ olduğundan H_0 reddedilir. Yani üç ilacın reaksiyon zamanları arasında fark vardır.

Farklı Grupların Belirlenmesi

$$d_c = z\sqrt{\frac{mk(k+1)}{6}} = 2{,}05\sqrt{\frac{(10)(3)(4)}{6}} = 9{,}17$$

$$z_{0,1/6} = z_{0,0167} = 2{,}05$$

$$|R_1 - R_2| = |13 - 19| = 6$$

$$|R_1 - R_3| = |13 - 28| = 15 > 9{,}17$$

$$|R_2 - R_3| = |19 - 28| = 9$$

Olduğundan 1. ve 3. ilaç arası anlamlı farklılık saptanmıştır. 1. ve 2. ilaç ile 2. ve 3. ilaç arası farklılık anlamlı bulunmamıştır.Yani 1. ilaç ile 3. ilacın reaksiyon zamanları farklıdır.Bu durumda A ilacı daha çabuk etki gösterir diyebiliriz.

```
x<-
c(10,10,11,8,7,15,14,10,9,10,10,15,15,12,12,15,12,14,9,14,15,20,16,10,13,15,1
8,17,12,16)
blok<-rep(1:10,3)
km<-c(rep("A",10),rep("B",10),rep("C",10))
friedman.test(y=datan$x,groups = datan$km,blocks = datan$blok)

##
##  Friedman rank sum test
##
## data:  datan$x, datan$km and datan$blok
## Friedman chi-squared = 22.026, df = 4, p-value = 0.0001981

if(!require(PMCMRplus)) {install.packages("PMCMRplus"); require(PMCMRplus)}
friedmanTest(data=datan,y=datan$x,groups = datan$km,blocks = datan$blok,dist
= "FDist")

##
##  Friedman rank sum test
##
## data:  y, groups and blocks
## Conover's F = 15.458, num df = 4, denom df = 28, p-value = 8.691e-07

frdAllPairsConoverTest(x,km,blok)

##
##  Pairwise comparisons using Conover's all-pairs test for a two-way
balanced complete block design

## data: y, groups and blocks
```

```
##   A      B
## B 0.3056 -
## C 0.0007 0.0704

##
## P value adjustment method: single-step
```

Örnek Tekrarlı ölçümlerde tek yönlü varyans analizi için verilen örneğin 11 öğrenci üzerinde yapıldığını düşünelim. Bu durumda Friedman testi için hazırlık tablosu aşağıdaki gibi olacaktır.

Öğrenci	İlk sınav Öncesi	İkinci sınav Öncesi	Üçüncü Sınav Öncesi	Sıra no R(1)	R(2)	R(3)
1	40	37	34	3	2	1
2	52	50	43	3	2	1
3	35	35	34	2.5	2.5	1
4	38	35	32	3	2	1
5	45	40	41	3	1	2
6	41	42	37	2	3	1
7	41	40	41	2.5	1	2.5
8	40	37	32	3	2	1
9	44	46	40	2	3	1
10	43	41	37	3	2	1
11	42	42	42	2	2	2

1. Adım Hipotezler kurulur.

H_0: Üç sınav öncesindeki durumluk kaygı puanları arasında fark yoktur (ya da durumluk kaygı puanları dönemlere göre değişmemiştir.)

H_1: Üç sınav öncesindeki durumluk kaygı puanları arasında fark vardır (ya da durumluk kaygı puanları dönemlere göre değişmiştir.)

2. Adım Test istatistiği hesaplanır.

Friedman testi için test istatistiği:

$$\chi_R^2 = \frac{12}{nk(k+1)}\left[\sum_{j=1}^{k} R_j^2\right] - \left[3n(k+1)\right]$$

n: Satır sayısı

k: Grup (sütun sayısı)

R_j: Her bir gruba (sütuna) ilişkin sıra numaraları toplamı

$$\chi_R^2 = \frac{12}{11\times 3\times (3+1)}\left[29^2 + (22{,}5)^2 + (14{,}5)^2\right] - \left[3\times 11\times (3+1)\right]$$

χ_R^2 **=11,105**

3. Adım Kritik tablo değeri bulunur.

$$\chi_{k-1,\alpha}^2 = \chi_{3-1,0.05}^2 = 5.9914$$

4. Adım Karar verilir.

$FR = 11.105 > 5.9914 = \chi_{2,0.05}^2$ olduğundan H_0 hipotezi reddedilir. Üç sınav öncesindeki durumluk kaygı puanları arasında fark vardır (ya da durumluk kaygı puanları dönemlere göre değişmiştir.)

```
x<-
c(40,52,35,38,45,41,41,40,44,43,42,37,50,35,35,40,42,40,37,46,41,42,34,43,34,
32,41,37,41,32,40,37,42)
blok<-rep(1:11,3)
km<-c(rep("A",11),rep("B",11),rep("C",11))
(datan<-data.frame(blok,km,x))

##    blok km  x
## 1     1  A 40
## 2     2  A 52
## 3     3  A 35
## 4     4  A 38
## 5     5  A 45
## 6     6  A 41
## 7     7  A 41
## 8     8  A 40
## 9     9  A 44
## 10   10  A 43
## 11   11  A 42
## 12    1  B 37
## 13    2  B 50
## 14    3  B 35
## 15    4  B 35
```

```
## 16    5  B 40
## 17    6  B 42
## 18    7  B 40
## 19    8  B 37
## 20    9  B 46
## 21   10  B 41
## 22   11  B 42
## 23    1  C 34
## 24    2  C 43
## 25    3  C 34
## 26    4  C 32
## 27    5  C 41
## 28    6  C 37
## 29    7  C 41
## 30    8  C 32
## 31    9  C 40
## 32   10  C 37
## 33   11  C 42
```

```
friedman.test(y=datan$x,groups = datan$km,blocks = datan$blok)
```

```
##
##  Friedman rank sum test
##
## data:  datan$x, datan$km and datan$blok
## Friedman chi-squared = 11.105, df = 2, p-value = 0.003877
```

```
if(!require(PMCMRplus)) {install.packages("PMCMRplus"); require(PMCMRplus)}
friedmanTest(data=datan,y=datan$x,groups = datan$km,blocks = datan$blok)
```

```
##
##  Friedman rank sum test
##
## data:  y, groups and blocks
## Friedman chi-squared = 11.105, df = 2, p-value = 0.003877
```

```
frdAllPairsConoverTest(x,km,blok)
```

```
##
##  Pairwise comparisons using Conover's all-pairs test for a two-way
balanced complete block design

## data: y, groups and blocks

##   A      B
## B 0.2894 -
## C 0.0023 0.1535

##
## P value adjustment method: single-step
```