



HACETTEPE ÜNİVERSİTESİ

SAĞLAM İSTATİSTİK -İST440

VİZE ÖDEVİ 2

2022-2023 BAHAR

19.04.2023

Sedanur Gülten- 21935965

İÇİNDEKİLER

1.BÖLÜM

Veri, Ortanca, Trimmed Ortalamalar,Winsorized Ortalama, Trimean, Broadened Ortalama.....2-6

2.BÖLÜM

Aykırı Değerli Yeni Veri,Ortanca, Trimmed Ortalamalar 2,Winsorized Ortalama 2, Trimean 2, Broadened Ortalama 27-10

3.Kaynakça.....11

Normal dağılımdan n=30 gözlemleri bir veri üretiniz ve bu veriye aykırı değer ekleyerek (aykırı değer sayısını 3 tane alabilirsiniz) yada veriyi bozarak ortanca, trimmed ortalama, winsorized ortalama, trimean, broadened ortalama konum kestiricilerini hesaplayınız.

Kesilmiş ve winsorezed ortalamayı hesaplariken %5, %10, %20 için hesaplayınız.

(Orijinal veri ve aykırı değerli veri kümesi için ayrı ayrı yapınız.)

Aşağıdaki veriler R programından normal dağılıma uyacak şekilde rastgele üretilmiştir. Üretmek için aşağıdaki kod kullanılmıştır. Küçükten büyüğe aşağıdaki gibi sıralanmıştır.

```
a<-random_data <- rnorm(30, mean = 0, sd = 1)
View(a)
```

	Sayılar	Yuvarlanmamış sayılar
1	-2,96	-2,960635808
2	-1,95	-1,946206915
3	-1,4	-1,402862866
4	-1,31	-0,306773546
5	-1,02	-1,019902257
6	-0,98	-0,980499317
7	-0,64	-0,643106252
8	-0,62	-0,622716104
9	-0,5	-0,501318864
10	-0,34	-0,336183295
11	-0,22	-0,21567354
12	-0,17	-0,174531824
13	-0,07	-0,06987551
14	-0,04	-0,041612
15	0	-0,001537027
16	0,05	0,046955362
17	0,13	0,133248242
18	0,15	0,146393524
19	0,33	0,331401046
20	0,49	0,486965659
21	0,49	0,491272447
22	0,56	0,557698059
23	0,6	0,600267899
24	0,72	0,717952715
25	0,89	0,886566936
26	1	0,996135863
27	1,06	1,062516677
28	1,14	1,137174034
29	1,35	1,351431359

30	1,57	1,565603754
	Toplam	Toplam
	-1,69	-0,711851549
	Ortalama	Ortalama
	-0,056333333	-0,023728385
	Ortanca	
	0,025	
	Birinci Çeyrek	
	-0,81	
	DördüncüÇeyrek	
	0,525	

Verimizde 30 Gözlem olduğu için, ortadan ikiye ayıran bir gözlem olmadığından 15. Ve 16. Gözlemin ortalamasıdır. $(0.05/2=0,025)$

Birinci çeyreklik 7.ve 8. Gözlemin ortalamasıdır. $((-0,98)+ (-0,62))/2= -0,81$

Dördüncü çeyreklik $(0,49+0,56)/2=0,525$

TRİMMED ORTALAMALAR

	T(0.05),n=26 silinen=4(2) trimmed	T(0.10),n=24 silinen=6 (3)	T(0.20),n=18 silinen=12 (6)
1			
2			
3	-1,4		
4	-1,31	-1,31	
5	-1,02	-1,02	
6	-0,98	-0,98	
7	-0,64	-0,64	-0,64
8	-0,62	-0,62	-0,62
9	-0,5	-0,5	-0,5
10	-0,34	-0,34	-0,34
11	-0,22	-0,22	-0,22
12	-0,17	-0,17	-0,17
13	-0,07	-0,07	-0,07
14	-0,04	-0,04	-0,04
15	0	0	0
16	0,05	0,05	0,05
17	0,13	0,13	0,13
18	0,15	0,15	0,15
19	0,33	0,33	0,33
20	0,49	0,49	0,49
21	0,49	0,49	0,49
22	0,56	0,56	0,56
23	0,6	0,6	0,6
24	0,72	0,72	0,72

25	0,89	0,89	
26	1	1	
27	1,06	1,06	
28	1,14		
29			
30			
	Toplam	Toplam	Toplam
	0,3	0,56	0,92

Trimmed kesilmiş ortalamalar hesabıdır. Yukarıda $n=30$ iken %5, %10 ve %20 için;

%5 için $n=30 \rightarrow 30 \cdot 0,05 = 1,5 \approx 2$ olduğundan, verinin sağ ve sol tarafı olmak üzere 4 gözlem çıkarılarak tekrar hesaplanmıştır. Burada $n=30$ sayısı, $n=26$ olarak değişir. Aynı işlemler diğer yüzdelikler için de yapılmıştır.

	İşlem	Ortalamalar
T(0,05)	0,3/26	0,011538462
T(0,10)	0,56/24	0,023333333
T(0,20)	0,92/18	0,051111111

WINSORIZED ORTALAMALAR

Aykırı veya etkili gözlemin etkisini azaltmak için kullanılabilir. Veriyi silmek yerine yüzdeliklerle belirlenen (0,05, 0,10 ,0,20) hesaplamada bulunan değiştirilmesi gereken verilerin sayısı belirlenir. Değişim için sağdan ve soldan bulunan sayı kadar değiştirme yapılır. Örneğin 2 bulduysak sağdan ikinci değerden sonra gelen değer ne ise (3.gözlemde bulunan değer) üçüncü gözleme kadar bu değer girilir, sol için de aynı durum tekrarlanır. Değişmeyen n sayısı ile yeni winsorized ortalama bulunur.

	W(0.05)-n=30 değiştirilen=4 (2)	W(0.10)-n=30 değiştirilen=6 (3)	W(0.20)-n=30 değiştirilen=12 (6)
1	-1,4	-1,31	-0,64
2	-1,4	-1,31	-0,64
3	-1,4	-1,31	-0,64
4	-1,31	-1,31	-0,64
5	-1,02	-1,02	-0,64
6	-0,98	-0,98	-0,64
7	-0,64	-0,64	-0,64
8	-0,62	-0,62	-0,62
9	-0,5	-0,5	-0,5
10	-0,34	-0,34	-0,34
11	-0,22	-0,22	-0,22
12	-0,17	-0,17	-0,17
13	-0,07	-0,07	-0,07
14	-0,04	-0,04	-0,04

15	0	0	0
16	0,05	0,05	0,05
17	0,13	0,13	0,13
18	0,15	0,15	0,15
19	0,33	0,33	0,33
20	0,49	0,49	0,49
21	0,49	0,49	0,49
22	0,56	0,56	0,56
23	0,6	0,6	0,6
24	0,72	0,72	0,72
25	0,89	0,89	0,72
26	1	1	0,72
27	1,06	1,06	0,72
28	1,14	1,06	0,72
29	1,14	1,06	0,72
30	1,14	1,06	0,72
	Toplam	Toplam	Toplam
	-0,22	-0,19	1,4

	İşlem	Ortalamalar
W(0,05)	-0,22/30	-0,007333333
W(0,10)	-0,19/30	-0,006333333
W(0,20)	1,4/30	0,046666667

TRİMEAN

Trimean için birinci çeyrek, ortanca ve dördüncü çeyreklik değerlerine ihtiyaç vardır.

Trimean= $\frac{1}{4}(F_1 + 2\mu + F_4)$ μ =ortancadır.

Trimean= $\frac{1}{4}((-0,81) + 2(0,025) + 0,525) = -0,1275$

BROADENED ORTALAMALAR

Genişletilmiş ortalama. Gözlemlerin yuvarlama, gruplama hatalarına karşı başarılıdır, aykırı değerlere karşı ortancanın dayanıklılığını korumaya çalışır.

n tek sayı olduğundan $n \geq 13$ olduğundan merkezdeki 6 sıralı istatistiğin ortalamasıdır. Ağırlıklar da aşağıdaki gibidir.

1.(1/10) 2.(1/5) 3.(1/5) 4.(1/5) 5.(1/5) 6.(1/10)

En baştaki, ilk veri setinden 13. ve 18. aralığındaki gözlemler ağırlıklandırılıp bölündüğünde, $1,62/6 = 0,27$ olarak genişletilmiş ortalama bulunur.

BÖLÜM 2

Aykırı Değer Eklenmiş Veri

Veriye 3 adet aykırı gözlem olabilecek değer ekledim.

	Aykırı Değerli Hali n=33		
1	-2,96		
2	-1,95		
3	-1,4		
4	-1,31		
5	-1,02		
6	-0,98		
7	-0,64		
8	-0,62		
9	-0,5		
10	-0,34		
11	-0,22		
12	-0,17		
13	-0,07		
14	-0,04		
15	0		
16	0,05		
17	0,13		
18	0,15		
19	0,33		
20	0,49		
21	0,49		
22	0,56		
23	0,6		
24	0,72		
25	0,89		
26	1		
27	1,06		
28	1,14		
29	1,35		
30	1,57		
31	57		
32	61		
33	70		
	Toplam		
	186,31		
	Ortalama		
	5,645757576		
	Ortanca	Birinci Çeyrek	Dördüncü Çeyrek
	0,13	-0,63	-0,56

Birinci Çeyrek = $(-0,64+(-0,62))/2$

Dördüncü Çeyrek = $(-0,62)+0,5=-1,12$ $-1,12/2=-0,56$

Toplamda 33 veri olan veri seti ile aynı işlemleri tekrar yapalım.

TRIMMED ORTALAMALAR-2

n=33 iken %5, %10 ve %20

%5 için $\rightarrow 33*0,05=1,65 \cong 2$

%10 için $\rightarrow 33*0,1=3,3 \cong 3$

%20 için $\rightarrow 33*0,2=6,6 \cong 7$

	T(0.05),n=29, silinen=4 (2), trimmed	T(0.10),n=27,silinen=6(3)	T(0.20),n=19,silinen=14(7)
1			
2			
3	-1,4		
4	-1,31	-1,31	
5	-1,02	-1,02	
6	-0,98	-0,98	
7	-0,64	-0,64	
8	-0,62	-0,62	-0,62
9	-0,5	-0,5	-0,5
10	-0,34	-0,34	-0,34
11	-0,22	-0,22	-0,22
12	-0,17	-0,17	-0,17
13	-0,07	-0,07	-0,07
14	-0,04	-0,04	-0,04
15	0	0	0
16	0,05	0,05	0,05
17	0,13	0,13	0,13
18	0,15	0,15	0,15
19	0,33	0,33	0,33
20	0,49	0,49	0,49
21	0,49	0,49	0,49
22	0,56	0,56	0,56
23	0,6	0,6	0,6
24	0,72	0,72	0,72
25	0,89	0,89	0,89
26	1	1	1
27	1,06	1,06	
28	1,14	1,14	
29	1,35	1,35	

30	1,57	1,57	
31	57		
32			
33			
	Toplam	Toplam	Toplam
	60,22	4,620	3,45

	İşlem	Ortalamalar
T(0,05)	60,22/29	2,076551724
T(0,10)	4,620/27	0,171111111
T(0,20)	3,45/19	0,181578947

WINSORIZED ORTALAMALAR-2

Yukarıda yüzdeliklerle hesaplanan (%5, %10, %20) değerlerce sıralanan verinin sağından ve solundan değiştirmeler yapıyoruz.

	W(0.05),n= 33,değiştirilen=4(2)	W(0.10)-n= 33 değiştirilen=6(3)	W(0.20),n= 33,değiştirilen=14(7)
1	-1,4	-1,31	-0,62
2	-1,4	-1,31	-0,62
3	-1,4	-1,31	-0,62
4	-1,31	-1,31	-0,62
5	-1,02	-1,02	-0,62
6	-0,98	-0,98	-0,62
7	-0,64	-0,64	-0,62
8	-0,62	-0,62	-0,62
9	-0,5	-0,5	-0,5
10	-0,34	-0,34	-0,34
11	-0,22	-0,22	-0,22
12	-0,17	-0,17	-0,17
13	-0,07	-0,07	-0,07
14	-0,04	-0,04	-0,04
15	0	0	0
16	0,05	0,05	0,05
17	0,13	0,13	0,13
18	0,15	0,15	0,15
19	0,33	0,33	0,33
20	0,49	0,49	0,49
21	0,49	0,49	0,49
22	0,56	0,56	0,56
23	0,6	0,6	0,6
24	0,72	0,72	0,72
25	0,89	0,89	0,89
26	1	1	1

27	1,06	1,06	1
28	1,14	1,14	1
29	1,35	1,35	1
30	1,57	1,57	1
31	57	1,57	1
32	57	1,57	1
33	57	1,57	1
	Toplam	Toplam	Toplam
	171,42	5,4	6,11
	Ortalama	Ortalama	Ortalama
	5,194545455	0,163636364	0,185151515

TRİMEAN-2

Trimean= $\frac{1}{4}(F_1 + 2\mu + F_4)$ μ =ortancadır.

Trimean= $\frac{1}{4}((-0,63) + 2(0,13) + (-0,56)) = -0,2325$ olarak bulunur.

BROADENED ORTALAMALAR-2

n=33, tek sayı olduğundan (n ≥ 13) merkezdeki 5 sıralı istatistiğin ortalaması olarak bulunur.

15 ve 19. Gözlem değerleri arasındaki sayıların toplamının 5'e bölümü -> $0,66/5 = 0,132$ olarak bulunur.

KAYNAKÇA

Hacettepe Üniversitesi, IST440-Sağlam istatistik ders notları, Meral Çetin