İstatistik-II Dersi

6.Bölüm: Parametrik Olmayan İstatistik

PARAMETRİK OLMAYAN İSTATİSTİK

İstatistikte yapılan testler parametrik ve parametrik olmayan testler olmak üzere iki gruba ayrılır. Birincisi yani parametrik testlerde değişkenlerin normal dağılması, varyansların homojen olması ve bazı durumlarda gözlem sayısının 30 ve üzeri olması gibi bazı varsayımlar vardır. Bu varsayımları test ettikten sonra eğer varsayımlar gerçekleşirse t-testi, Z-testi ya da ANOVA testleri gibi testler gerçekleştirilebilir. İkinci testler ise bu varsayımların geçerli olmadığı durumlarda kullanılabilen parametrik olmayan(non-parametric) testlerdir.

Parametrik olmayan istatistiksel yöntemler; anakütle hakkında, genellikle sayısı çok az olan varsayımlara bağlıdır. Bu testlerin kullanılabilmesi için anakütle dağılımının bilinmesi gerekmemektedir.

Parametrik olmayan istatistiksel yöntemler; sınıflayıcı ölçek (nominal) veya sıralayıcı ölçeklere (ordinal) göre elde edilen verilerde daha başarılı sonuçlar vermektedirler.

Gözlemlerden elde edilen veriler genellikle sıralandıktan veya işarete dönüştürüldükten sonra test edilir.

Parametrik Olmayan Yöntemlerin Avantajları

- Birçok parametrik olmayan istatistiksel teknikler en az varsayıma dayandıkları için yanlış uygulanma ihtimali daha azdır.
- Bazı parametrik olmayan yöntemler, alternatifleri olan parametrik yöntemlere göre daha kolay ve daha hızlı hesaplanabilir.
- Parametrik olmayan istatistiksel tekniklerin çoğu, çok fazla istatistik yapısı gerektirmediğinden, sosyal bilimlerde veya eğitim bilimlerinde kolayca uygulanmaktadır.
- Zayıf ölçekle oluşturulmuş veri türlerinde kolayca uygulanabilir.

Parametrik Olmayan Yöntemlerin Dezavantajları

- Parametrik olmayan yöntemler, parametrik yöntemlere göre hesaplama kolaylığından dolayı tercih edilebilirler. Fakat bu durumda, daha doğru sonuçlar verebilen parametrik yöntemler göz ardı edilebilir ve sonuçlar yanıltıcı olabilir.
- Örnek hacimlerinin geniş olması, uzun ve tutarsız sonuçlar verebilir.

Parametrik Olmayan Testlerin Sınıflandırılması

Parametrik olmayan testlerin sınıflandırılması, uygulandıkları örneklem türlerine göre aşağıda verilmiştir. Verilen bu sınıflama tüm parametrik olmayan testleri kapsamamakla beraber, en çok kullanılan testleri göstermektedir.

Uyum İyiliği Testleri

- Ki-Kare uyum iyiliği testi
- Kolmogorov-Smirnov uyum iyiliği testi
- Normallik için Lilliefors testi
- Normallik için Shapiro-Wilks testi

Tek örneklem testleri

- Binom testi
- Tek örneklem işaret testi
- Wilcoxon işaret testi
- Run testi

İki bağımsız örneklem için testler

- Medyan Testi
- Mann-Whitney U Testi
- Wald-Wolfowitz Dizi Parçaları Testi
- Kolmogorov-Smirnov Testi
- Fischer Tam Olasılık Testi
- Mood Testi
- Siegel-Tukey Testi

İki bağımlı örneklem için testler

- İşaret testi
- Wilcoxon işaretli sıra sayıları testi
- McNemar testi
- Kappa testi

İkiden fazla bağımsız örneklem için testler

- Medyan testi
- Kruskal-Wallis H testi

İkiden fazla bağımlı örneklem için testler

- Friedman S testi

- Cochran Q testi

İlişki katsayıları

- Spearman sıra korelasyonu katsayısı
- Kendall τ ilişki katsayısı
- Goodman ve Kruskal Gamma Korelasyon Katsayısı

Tek örneklem için Testler

Tek örneklem için hipotez testlerinin, parametrik olmayan karşılıklarının kullanıldığı testlerdir. Normal dağılım göstermeyen veya yeterince küçük örneklem hacimlerine sahip verileri test etmek için kullanılır. Bu testler ile birlikte, anakütleye ait medyan parametresi üzerinde tanımlanan hipotezlerin test edilerek karşılaştırılması yapılmaktadır. Burada en çok kullanılan tek örneklem testlerinden "tek örneklem işaret testi" tanıtılacaktır.

Tek örneklem işaret testi

Binom testinin özel bir durumu olan bu test parametrik olmayan testlerin en eskisi olup, ilk olarak 1710 yılında John Arbuthnot tarafından kullanılmıştır. Tek örneklem işaret testinde ilgilenilen değişken sürekli olmalıdır. Bunun yanısıra veriler en az sıralama düzeyinde ölçülmüş olmalı ve örneklem medyanı bilinmeyen bir anakütleden çekilmiş olmalıdır.

i) *M*: Örneğin çekildiği kitleye ait bilinmeyen medya parametresi ve M_0: herhangi bir reel sayı olmak üzere:

a) $H_0: M = M_0$	b) $H_0: M = M_0$	c) $H_0: M = M_0$
$H_0: M < M_0$	$H_0: M > M_0$	$H_0: M \neq M_0$

ii) Anakütleden *n* birimlik örnek çekilir.

iii) H_0 hipotezi dikkate alınarak $i=1,2,\ldots,n$ için $D_i = x_i - M_0$ farkları dikkate alınarak, n birimlik örneklem (+) veya (-) işaretler dizisine dönüştürülür. Eğer $D_i = 0$ sonucunu veren gözlem değerleri varsa, bu değerler işlem dışı tutularak hesaplamaya devam edilir. Yeni örneklem hacmi, işlem dışı tutulan gözlemlerden sonra kalan örneklem sayısı olmaktadır.

iv) Test istatistiği hesaplanarak, test istatistiğinin örnekleme dağılımı belirlenir. *k:* Test istatistiği olmak üzere; ya (+) işaretlilerin ya da (-) işaretlilerin sayısını ifade etmektedir. Bu işaretlerin hangisinin seçileceği H_1 hipotezine göre belirlenmektedir. Daha az sayıda çıkmasını beklediğimiz işaretlilerin sayısı test istatistiği olarak alınır. Örneklemler bağımsız olarak seçildiğinden ve + veya – gibi iki sonucu bulunduğundan, *k* istatistiği Binom dağılımına yakınsayacaktır.

Test istatistiği ve test istatistiğinin örnekleme dağılımı belirlenir. İşaret testinde test istatistiği;

$k \sim Binom\left(n,\ \theta = \frac{1}{2}\right)$ olup, olasılık fonksiyonu $f(k) = \binom{n}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k}$, k =0,1, 2, ..,n dir.

k istatistiğinin hesaplanan değeri k_h olmak üzere karar verme kuralı:

H_1 Hipotezi	**Karar Kuralı**
$M \neq M_0$	k_h=(+)veya(-)işaretlilerin minimumu olmak üzere
$Pr(k \leq k_h) = p$ olmak üzere;	$p < \alpha/2$ ise H_0 red edilir.
$M > M_0$	k_h (-) işaretlilerin sayısı olmak üzere
$Pr(k \leq k_h) = p$ olmak üzere;	$p < \alpha$ ise H_0 red edilir.
$M < M_0$	k_h (+) işaretlilerin sayısı olmak üzere
$Pr(k \leq k_h) = p$ olmak üzere;	$p < \alpha$ ise H_0 red edilir.

Örnek-1: X-ray tekniği ile tedavi gören hastaların iyileşme sürelerine ait medyan parametresinin 30 günden fazla olup olmadığı araştırılmak isteniyor. Bunun için 8 hastanın iyileşme süreleri ölçülmüş ve aşağıda verilmiştir. %5 anlam düzeyinde işaret testi ile iddiayı araştırınız.

İyileşme süreleri: 32 65 28 33 61 46 29 51

Çözüm:

Hipotezler: $H_0: M = 30$ $H_1: M > 30$

İyileşme süresi	$\boldsymbol{D_i = x_i - M_0}$	**İşaret**
32	32-30=2	+
65	65-30=35	+
28	28-30=-2	-
33	33-30=3	+
61	61-30=31	+
46	46-30=16	+
29	29-30=-1	-
51	51-30=21	+

Test istatistiği H_1 hipotezine göre k_h =(-) işaretlilerin sayısı işaretlilerin sayısı olacaktır.

$$k \sim Binom(n = 8, p = \frac{1}{2})$$

k' nın olasılık yoğunluk fonksiyonu ise şu şekilde olacaktır:

$$f(k) = \binom{8}{k}\left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{8-k}$$

(-) 2 tane olduğundan

$$\Pr(k \le 2) = \sum_{k=0}^{2} \binom{8}{k} \left(\frac{1}{2}\right)^{k} \left(\frac{1}{2}\right)^{8-k}$$

$$= \binom{8}{0} \left(\frac{1}{2}\right)^{0} \left(\frac{1}{2}\right)^{8-0} + \binom{8}{1} \left(\frac{1}{2}\right)^{1} \left(\frac{1}{2}\right)^{8-1} + \binom{8}{2} \left(\frac{1}{2}\right)^{2} \left(\frac{1}{2}\right)^{8-2}$$

$$= \frac{1}{2^8}\left[\binom{8}{0} + \binom{8}{1} + \binom{8}{2}\right] = 0.14 = p$$

$p = 0.14 > \alpha = 0.05$ olduğundan H_0 hipotezi red edilemez. Medyan parametresinin %95 güvenle 30'dan büyük olduğunu söyleyecek yeterli kanıt yoktur.

Örnek-2: Her hafta oynanan Türkiye 1. Ligi maçlarında berabere biten maç sayısı ortalama 3'tür. Aşağıda verilen ilk 10 haftalık beraberlik sayıları yardımıyla, 11. Haftada 3 beraberlik olacağı konusunda %5 anlam düzeyinde karar veriniz.

Hafta	1	2	3	4	5	6	7	8	9	10
Beraberlik sayısı	3	1	1	2	6	1	4	2	3	4

Büyük Hacimli Örneklerde İşaret Testi

İşaret testinde örnek hacminin büyük olması özellikle n > 20 iken test istatistiğinin örnekleme dağılımı kullanılarak olasılık hesaplamalarında işlem uzunluğu bakımından sıkıntılar ortaya çıkabilmektedir. Böyle durumlarda Binom dağılımının normal dağılıma yaklaşımından yararlanarak test istatistiğinin örnekleme dağılımını Binom dağılımından normal dağılıma taşımak mümkündür. Büyük örnek durumunda işaret testi için test istatistiği, Binom dağılımının kesikliği ve normal dağılımın sürekli olması nedeniyle uygulanacak olan süreklilik düzeltmesi ile birlikte

$$\boxed{Z = \frac{\left(k \pm \frac{1}{2}\right) - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \sim N(0\,,1)}$$

şeklinde tanımlanır. Burada düzeltme terimi (k ± 1/ 2)'dır.

İki bağımsız örneklem için testler

İki bağımsız örneklem karşılaştırmasında kullanılan t ve F testleri en çok kullanılan parametrik testlerdir.(İki ortalamanın eşitliği için kullanılan t-testi ve iki varyansın eşitliği için kullanılan F-testi) Bu testler, anakütleden çekilen örneklem dağılımlarının normal olması varsayımına dayanmaktadır. Karşılaştırılan iki örnekten en az birinin dağılımı normal değilse veya en az birinin örneklem hacmi oldukça küçük ise parametrik testler güvenilir sonuç vermeyecektir. Bunların yerine kullanılan parametrik olmayan testlerden en çok kullanılan Mann-Whitney U testidir.

Mann-Whitney U testi

Medyan testi gibi iki bağımsız grubu medyan parametreleri yönünden karşılaştırmada kullanılan bir parametrik olmayan test tekniğidir.

Varsayımları i) İlgilenilen değişken sürekli olmalıdır. ii) İlgilenilen değişkenin ölçme düzeyi en az sıralama olmalıdır. iii) Veriler iki bağımsız gruptan rastgele ve birbirinden bağımsız olarak çekilmelidir.

n_1 ve n_2 hacimli bağımsız iki örneğin sıralama ya da aralıklı ölçekle elde edilmiş verilerinin aynı medyanlı populasyondan alınmış rasgele örnekler olup olmadığını test etmek için Mann-Whitney U testi kullanılır.

Mann-Whitney U testinde test edilen hipotezler aşağıdaki gibidir.

H_0 : n_1 ve n_2 veri setleri aynı medyanlı dağılıma sahiptir.

H_1 : Örnekler farklı medyanlı dağılımların örnekleridir.

H_1 : n_1 veri setinin gözlemlerinin yarısından fazlası diğer setten farklıdır.

H_1 :$P(a>b)\neq 1/2$

U testi uygulamak için n_1 ve n_2 hacimli iki örnek bir tek dizi (genel dizi) haline getirilir ve dizideki gözlemlerin sıralama puanları bulunur.

Küçükten büyüğe doğru her gözlemin genel dizide kaçıncı sırada yer aldığı belirlenir.

Sıralı Dizi yeniden ele alınır ve her verinin hangi örneğe ait olduğu dikkate alınarak sıralama puanları örneklere göre toplanır.

1. örneğe ait gözlemlerin sıralama puanları toplamı R1 ,

2. örneğe ait olanları toplamı R2 bulunur.

Birim sayıları ve toplam sıralama puanlarından yararlanarak U1 ve U2 test istatistikleri hesaplanır.

U_1 ve U_2 test istatistikleri aşağıdaki gibi hesaplanır.

$$U_1 = n_1 * n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 * n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

U_1 ve U_2 ’den küçük olanı U test istatistiği olarak alınır.

Eğer $U_1 \leq U_2$ **ise** $U = U_1$,
$U_1 > U_2$ **ise** $U = U_2$ **alınır.**

$n_1 > 20$ ve $n_2 > 20$ ise U’nun önemliliği normal yaklaşımla bulunur. Bunun için U’nun ortalama ve standart sapması bulunur ve z test istatistiği aşağıdaki gibi hesaplanır.

$$\mu_U = (n_1)(n_2)/2$$

$$\sigma_U = \sqrt{\frac{(n_1)(n_2)(n_1+n_2+1)}{12}}$$

$$z = \frac{U-\mu_U}{\sigma_U} = \frac{U-(n_1)(n_2)/2}{\sqrt{\frac{(n_1)(n_2)(n_1+n_2+1)}{12}}}$$

Örnek-3: Rasgele seçilen 20 X hastası iki gruba ayrılmıştır. 1. Gruptaki 12 hasta A yöntemi ile 2. Gruptaki 8 hasta da B yöntemi ile tedavi edilmişlerdir. Bu hastalara iyileşme durumlarına göre verilen puanlar aşağıdaki tabloda verilmiştir. A ve B yöntemi arasında fark var mıdır? Test ederek tartışınız.

A Grubu	B Grubu
30	37
20	80
40	82
60	59
41	38
50	91
90	88
15	68
47	
56	
12	
51	

$$H_0: M_1 = M_2$$

$$H_1: M_1 \neq M_2$$

A Grubu	Sıralama Puanları (R_1)	B Grubu	Sıralama Puanları (R_2)
30	4	37	5
20	3	80	16
40	7	82	17
60	14	59	13
41	8	38	6
50	10	91	20
90	19	88	18
15	2	68	15
47	9		
56	12		
12	1		
51	11		
Toplam	R_1=100		R_2=110

$$U_1 = n_1 * n_2 + \frac{n_1(n_1+1)}{2} - R_1 = 12*8 + \frac{12(12+1)}{2} - 100 = 74$$

$$U_2 = n_1 * n_2 + \frac{n_2(n_2+1)}{2} - R_2 = 12*8 + \frac{8(8+1)}{2} - 110 = 22$$

$$z = \frac{22-(12)(8)/2}{\sqrt{\frac{(12)(8)(12+8+1)}{12}}} = -2.01$$

U=22, Z=- 2.01, P<0.05*. A ve B tedavi yöntemleri arasında fark vardır.

Örnek-4: İstatistik bölümünde okuyan öğrencilerden 20 tanesi rastgele seçilmiş ve öğrenciler iki gruba ayrılmışlardır. Birinci gruptaki öğrenciler bilgisayar destekli dil eğitimi alırken, ikinci gruptaki öğrenciler yabancı bir İngilizce hocası ile dil eğitimi almışlardır. Dönem sonunda her iki gruba ortak İngilizce sınavı yapılmıştır. Aşağıda verilen sonuçlar için birinci gruptaki öğrencilerin başarı puanlarının ikinci gruptaki öğrencilerin başarı puanlarından daha az olduğu iddiasını %5 önem düzeyinde test ediniz.

1.grup	20	40	30	60	45	50	90	15	47	56	12	51
2.grup	37	80	82	59	38	91	88	68				

Örnek-5: AB üye ülkeleri ile AB üyesi olmayan ülkelerin Eurovision yarışmasından aldıkları puanlar aşağıda verilmiştir.%5 önem seviyesinde üye ülkelerin daha çok puan aldıklarını söyleyebilir misiniz?

Üye ülkeler	Alınan puan	Üye olmayan ülkeler	Alınan puan
İsveç	372	Arnavutluk	146
Almanya	110	Moldova	81
Romanya	71	Azerbaycan	150
Sırbistan	214	Makedonya	71
İtalya	101	Türkiye	112
Rusya	259	Estonya	120
İspanya	97	Bosna Hersek	55
Litvanya	70		
Ukrayna	65		
Yunanistan	64		

İki bağımlı örneklem için testler

İki bağımlı örnek verileri genel olarak iki farklı şekilde derlenir. Birinci yöntemde veriler, işlem öncesi-işlem sonrası veri toplanması ile düzenlenerek toplanır. İkinci yöntemde ise bağımlı değişkenler eşleştirilerek iki gruba ayrılır ve hesaplama yapılır. Bu alt başlıkta Wilcoxon işaretli sıra sayıları testi kısaca tanıtılacaktır.

Wilcoxon İşaretli Sıra Sayıları Testi (Wilcoxon t testi)

Anakütleden çekilen ve birbirine bağımlı iki örneklemin aynı dağılım gösterip göstermediğini belirlemede kullanılan bir testtir. Eşleştirilmiş t-testinin parametrik olmayan alternatifidir. Wilcoxon testi, iki örneklemi tek bir örnekleme indirgeyerek çözümleme yapar.

Testin algoritması şu şekildedir:

i) İki eşleştirilmiş örneklemin farkı $D_i = X_1 - X_2$ formülü ile alınır. Bu farkların mutlak değerleri hesaplanır.

ii) Farklar küçükten büyüğe doğru sıralanır. Farkların eşit olduğu durumlarda rankların ortalaması alınır.

iii) Mutlak değerleri alınmış ranklar pozitif $(X_1 > X_2)$ veya negatif $(X_1 < X_2)$ olarak işaretlenir. $(X_1 = X_2)$ olan çiftler analizden çıkarılır.

iv) Pozitif veya negatif rankların toplamı hesaplanır.

v) Hesaplanan Wilcoxon istatistiği, Wilcoxon tablo değeri ile karşılaştırılarak karar verilir.

Kurulan hipotezin tek yönlü veya çift yönlü olmasına göre, test istatistiği farklılık gösterecektir. Hipotezin kurulmasına göre verilecek kararlar aşağıda özetlenmiştir. $(W_C: Wilcoxon\ Test\ İstatistiği\ ve\ W_T: Wilcoxon\ Tablo\ Değeridir)$

Çift Yönlü Hipotez	**Karar**
$H_0: X_A = X_B$ $H_1: X_A \neq X_B$	Test istatistiği min$(\sum R^+, \sum R^-)$ olarak alınır. $W_C \leq W_T = W_{\alpha/2}$ ise H_0 hipotezi red edilir.
Tek Yönlü Hipotez	**Karar**
$H_0: X_A = X_B$ $H_1: X_A > X_B$	Test istatistiği $(\sum R^-)$olarak alınır. $W_C \leq W_T = W_\alpha$ ise H_0 hipotezi red edilir.
veya	
$H_0: X_A = X_B$ $H_1: X_A < X_B$	Test istatistiği $(\sum R^+)$olarak alınır. $W_C \leq W_T = W_\alpha$ ise H_0 hipotezi red edilir.

Yararlanılan Kaynaklar

1. Newbold, P. (2005). İşletme ve İktisat İçin İstatistik. (Ü. Şenesen, Çev.). İstanbul: Literatür Yayıncılık.

2. Erilli, N.A. (2018). İstatistik-2. Ankara: Seçkin Yayıncılık.

3. Spiegel, M. R., & Stephens, L. J. (1988). İstatistik. Çev. Erdal Öney ve Nahit Töre. Ankara: AÜ Basımevi.

4. McClave, J. T., & Sincich, T. (2006). Statistics (No. QA 276.12. M33 2006).

5. Cengiz, M. A. & Terzi, Y. (2020) Hipotez Testleri Ders Notları