



Araştırma Yöntemleri

Parametrik Olmayan Testler

Plan

- Son iki derste konuşulan konular;
 - Hipotez testleri, Tür 1-Tür 2 hataları, etki boyutu ve istatistik testlerin gücü ile
 - parametrik testlerdi (t-, ANOVA, korelasyon ve regresyon).
- Bu derste **parametrik olmayan** testler incelenecektir.

Not: Parametrik olmayan testler için Field, 2005, Bölüm 13 ve Field ve Hole, 2008 Bölüm 7'den yararlanılmıştır



Parametrik Olmayan Testler

Burada artık;

- Verilerin normal dağılmış olması gerekmiyor.
- Veriler sınıflama ya da sıralama ölçme düzeyinde toplanmış olmalı.
- Ya da eşit aralıklı/oranlı veriler sıralandıktan sonra test yapılacaktır.
- **Parametrik testlerden daha az güçlü**
- **Tür 2 hatası yapma olasılığı daha yüksek**
(gerçekte gruplar arasında fark varken fark olmadığını kabul etme olasılığı daha yüksek)



Parametrik Olmayan Testler Hangileri?

- Binom testi
- Ki- kare testi
- Mann-Whitney testi
- Wilcoxon işaretli sıralar toplamı ve işaret testi
- Kruskal Wallis testi
- Friedman'ın ANOVA testi
- Parametrik olmayan korelasyon testi

Veriler

- Aksi belirtilmedikçe testler için kullanılan veri dosyası: **hsb2turkce.sav**
- Veri dosyası 200 lise öğrencisine ait cinsiyet, ırk, sosyo-ekonomik statü, okul türü gibi demografik bilgileri ve öğrencilerin okuma, yazma, matematik, fen ve sosyal bilimler derslerinden aldıkları standart puanları içermektedir.

Veri Dosyasındaki Değişkenlerin Tanımları

No: Denek numarası

Cinsiyet: 0=erkek, 1=kadın (sınıflama)

Irk: 1=Latin, 2=Asyalı, 3=Siyah, 4=Beyaz (sınıflama)

Sosyo-ekonomik statü (sed): 1=düşük, 2=orta, 3=yüksek (sıralama)

Okul türü: 1=devlet, 2=özel (sınıflama)

Program türü: 1=genel, 2=akademik, 3=mesleki (sınıflama)

Okuma puanı (*okumanot*): Bu dersten aldığı not (oranlı)

Yazma puanı (*yazmanot*): Bu dersten aldığı not (oranlı)

Matematik puanı (*matnot*): Bu dersten aldığı not (oranlı)

Fen puanı (*fennot*): Bu dersten aldığı not (oranlı)

Sosyal bilimler puanı (*sosnot*): Bu dersten aldığı not (oranlı)

Binom Testi Örneđi

- Tek örneklemlili binom testi sınıflama ölçme düzeyinde toplanmış veriler için kullanılır. Bağımlı deđişken hakkındaki veriler iki düzeylidir (ör., Erkek-Kadın)
- Verilerin öngörülen bir yüzdeden farklı olup olmadığını test etmek için kullanılır.
- Örneđin «öğrencilerin cinsiyete göre dağılımı %50'den (yani 0,5) farklı mıdır?», sorusuna yanıt aransın.
- Araştırma hipotezi (H_1): “Öğrencilerin cinsiyete göre dağılımı eşit değildir.” ($H_1: \mu \neq \mu_0$) (çift kuyruk testi)

Seenek Hipotezler

- Arařtırma hipotezi byktr/kktr diye de kurulabilir. O zaman tek kuyruk (bykse sol, kkse saė) testi yapılır.
- H_1 : “Kız ėrencilerin oranı %50’den daha yksektir.” $H_1 : \mu < \mu_0$ (sol kuyruk testi)
- H_0 : “Kız ėrencilerin oranı %50’den daha dřktr.” $H_0 : \mu > \mu_0$

Binom Testi - PASW

Menüden;

- Analyze -> Nonparametric Tests-> Legacy Dialogs -> Binomial'i seçin.
- Test değişkenleri olarak Cinsiyet'i seçin.
- Test oranı olarak 0,5 girin.
- OK seçeneğine basın

Binom Testi Sonucu

Binom testi

Binomial Test

	Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (2-tailed)
ogrencinin cinsiyeti	Group 1 erkek	91	,46	,50	,229 ^a
	Group 2 kadin	109	,55		
	Total	200	1,00		

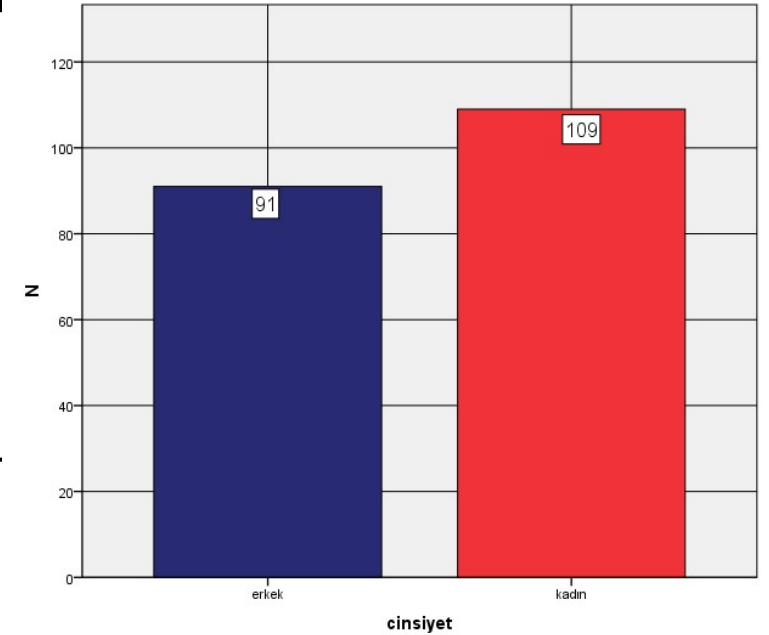
a. Based on Z Approximation.

- Deneklerin %46'sı (91) erkek, %55'i kız (109). Toplamın %101 olması yuvarlama hatasından kaynaklanıyor.
- Bu yüzdeler test değerinden (%50) farklı.
- Ama bu fark istatistiksel açıdan anlamlı değil ($p = 0,229$).
- Yani şansa bağlı olarak böyle bir oran elde edilebilir.
- Başka bir deyişle, cinsiyete göre dağılım %50'ye eşittir.

Boş hipotez kabul edilir.

Binom Testinin Yorumu

- Deneklerin cinsiyete göre dağılım yüzdeleri (%46 erkek, %55 kız) arasında anlamlı bir fark yoktur ($p = 0,229$). Başka bir deyişle, örneklemdaki erkeklerin oranı %50 değerinden anlamlı bir biçimde farklı değildir.
- Basit tanımlayıcı istatistikler için histogram vermeye gerek yoktur. Yorum yandaki gibi abartılı bir şekilde desteklenirse okuyucuya daha çok bilgi sunulmuş olmaz.
- **Tablolar/şekiller metindeki bilgilere yeni bilgi ekliyorsa kullanılmalıdır.**



Şekil 1. Deneklerin cinsiyete göre dağılımı (N=200)

Ki- kare Testleri

- Ki- kare (χ^2) testi sınıflama ölçüm düzeyinde toplanan veriler için kullanılır.
- Değişkenlerin gözlenen sıklık değerleriyle beklenen sıklık değerleri arasındaki farkı test eder.
- χ^2 test sonucu ikisi arasında istatistiksel açıdan anlamlı bir fark olduğunu gösterirse bu farkın şans eseri olmadığı sonucuna varılır.
- İki tür χ^2 testi vardır:
 - χ^2 uyum iyiliği testi
 - χ^2 ilişki testi

Ki- kare Uyum İyiliği Testi

- Bir bağımsız değişken için gözlenen sıklık değerlerinin beklenenlerden farkını test eder.
- Örneğin öğrencilerin sosyo-ekonomik durumlarının evrendeki %30 düşük, %50 orta, %20 yüksek oranlarına uygun olup olmadığını test edelim.
- Araştırma hipotezi (H_1): “Öğrencilerin sosyo-ekonomik durumlarına göre dağılımı %30 düşük, %50 orta, %20 yüksek şeklinde değildir” ($H_1 : \mu \neq \mu_0$) (çift kuyruk testi).
- Veya (H_1): “Sosyo-ekonomik durumu düşük olan öğrencilerin oranı %30’un altındadır” ($H_1 : \mu < \mu_0$) (sol kuyruk testi)

Ki- kare Uyum İyiliği Testi - PASW

Menüden;

- Analyze -> Nonparametric Tests -> Legacy Dialogs -> Chi-square'i seçin.
- Test değişkeni olarak öğrencinin sosyoekonomik durumunu seçin.
- Beklenen değerler olarak Values kısmına sırasıyla 30, 50, 20 girin.
- OK'e tıklayın.

Ki- kare Uyum İyiliği Testi Sonucu

sosyo-ekonomik durum

	Gözlenen (N)	Beklenen (N)	Sapma
düşük	47	60,0	-13,0
orta	95	100,0	-5,0
yüksek	58	40,0	18,0
Toplam	200		

Test Statistics

	sosyo-ekonomik durum
Chi-square	11,167 ^a
df	2
Asymp. Sig.	,004

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 40,0.

- İlk tablo gözlenen ve beklenen değerleri ve ikisi arasındaki sapma miktarını verir.
- 2. tablo χ^2 (ki kare) değerini (11,167), serbestlik derecesini (SD=2) ve p değerini ($p = 0,004$) verir.
- χ^2 testinde sıklık değeri 5'ten az olan ya da sıfır olan göz olmazsa daha isabetli sonuç verir. Bu koşul ihlal edilirse Fisher kesin testi kullanılır. (Bu örnekte böyle bir sorun yok)

Yorum

- Bu sonuç örneklemdaki öğrencilerin sosyo ekonomik durumlarına göre dağılımının hipotezde öngörülen değerlerden farklı olduğunu göstermektedir.
- Ekonomik durumu düşük olanların oranı beklenen %30'dan daha düşük iken, yüksek olanların oranı beklenenden (%20) daha yüksektir.
- Ki- kare istatistiği ve p değeri de bunu gösteriyor (ki-kare=11,167; SD=2; $p=0,004$) (*SD sosyo-ekonomik sınıf sayısının bir eksigidir*).

Boş hipotez reddedilir.

Rapor Etme

APA stiline göre bu bulgu şöyle rapor edilir:

- “Öğrencilerin sosyo-ekonomik durumlarına (düşük, orta, yüksek) göre dağılımı evrendeki dağılımdan -beklenen dağılım- anlamlı düzeyde farklıdır ($\chi^2_{(2)} = 11,167, p = 0,004$). Örneklemdeki sosyo-ekonomik durumu yüksek olan öğrencilerin oranı beklenenden daha yüksek, düşük olanların oranı beklenenden daha düşüktür.”

Ki- kare İlişki Testi

- Ki- kare testi iki bağımsız sınıflama değişkeni arasında ilişki olup olmadığını test etmek için kullanılır.
- Örneğin «öğrencilerin cinsiyetiyle mezun oldukları lise türü (genel/anadolu/meslek) arasında bir ilişki olup olmadığı» sorusuna yanıt aransın.
- Araştırma hipotezi (H_1): “Öğrencilerin cinsiyetiyle mezun oldukları lise türü arasında bir ilişki vardır.” (çift kuyruk testi)

Ki- kare İlişki Testi - PASW

Menüden;

- Analyze -> Descriptive statistics -> Crosstabs'i seçin
- Satıra lise türünü, sütuna cinsiyeti yerleştirin.
- Statistics seçeneğine tıklayarak Chi square'i işaretleyin.
- Cells seçeneğine tıklayarak Observed ve Expected'i işaretleyin.
- OK'e tıklayın.

Ki- kare Testi Sonucu

program türü * cinsiyet çapraz tablosu

			cinsiyet		Toplam
			erkek	kadın	
program türü	genel	Gözlenen	21	24	45
		Beklenen	20,5	24,5	45,0
	anadolu	Gözlenen	47	58	105
		Beklenen	47,8	57,2	105,0
	meslek	Gözlenen	23	27	50
		Beklenen	22,8	27,3	50,0
Toplam		Gözlenen	91	109	200
		Beklenen	91,0	109,0	200,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	,053 ^a	2	,974
Likelihood Ratio	,053	2	,974
Linear-by-Linear Association	,003	1	,955
N of Valid Cases	200		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 20,48.

- Cinsiyet ile mezun olunan lise türü çapraz tablosunda gözlenen ve beklenen değerler birbirine çok yakın
- Nitekim χ^2 (ki kare) değeri (0,053) çok küçük ve anlamlı değil (SD=2, $p = 0,974$) (Sıklık değeri satır (3) ve sütün (2) sayılarından birer çıkarılıp birbiriyle çarpılarak elde edilir

Boş hipotez kabul edilir.

Yorum

APA stiline göre bu bulgu şöyle rapor edilir:

- “Öğrencilerin cinsiyetiyle mezun oldukları lise türü arasında istatistiksel açıdan anlamlı bir ilişki yoktur ($\chi^2_{(2)} = 0,053$, $p = 0,974$).”

Mann-Whitney Testi

- Bağımsız örneklem t- testinin parametrik olmayan karşılığıdır. Bağımlı değişkenin sıralı olduğu varsayılır.
- **Mann-Whitney** testi bağımsız örneklem t- testine çok benzer.
- Araştırma hipotezi (H_1): “Erkek ve kız öğrencilerin yazma puanlarının ortalaması birbirinden farklıdır.” $H_1: \mu \neq \mu_0$ (çift kuyruk testi).
- Önce bağımlı değişkenin (yazma puanı) normal dağılıma uygun olup olmadığı **Kolmogorov-Smirnov** (K-S) testi ile test edilir.

Önce Kolmogorov-Smirnov Normallik Testi

Menüden;

- Analyze -> Descriptive Statistics-> Explore'u seçin.
- Bağımlı değişkene yazma puanını atayın.
- Bağımsız değişkene (Factor List) Cinsiyeti atayın.
- Both seçeneği işaretli (İstatistikler ve Grafik için).
- Statistics sekmesinde Descriptives'in ve %95 güven aralığı seçili (değiştirilebilir).
- Plots (grafik) sekmesinden Normality plots with test seçeneğini işaretleyin.
- OK'e tıklayın.

K-S Normallik Testi Sonucu

cinsiyet		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
yazma notu	erkek	,124	91	,001	,950	91	,002
	kadın	,139	109	,000	,946	109	,000

a. Lilliefors Significance Correction

- K-S testine göre hem erkeklerin hem de kızların yazma notları Sig. değerleri 0,05'in altında olduğu için **normal dağılım koşulunu yerine getirmiyor**.
- Shapiro-Wilk testi de K-S benzeri bir normallik testi (bu test de benzer sonuç vermiş).

K-S Testi Sonucunu Rapor Etme

K-S testi sonucu D ile gösterilir ve APA stiline göre şöyle rapor edilir:

- “Hem erkeklerin ($D(91)=0,124, p=0,001$) hem de kızların ($D(109)=0,139, p=0,000$) yazma notları **normal dağılıma uygun değildir**”.
- Bu durumda verilere parametrik olmayan test uygulanır.
- Daha önce niçin parametrik test uyguladığımız sorusunun yanıtı testin sonunda.

Mann-Whitney Testi - PASW

Menüden:

- Analyze -> Nonparametric Tests-> Legacy Dialogs -> 2 Independent Samples'i seçin
- Test değişkeni olarak yazma puanını seçin.
- Grup değişkeni olarak cinsiyeti seçin.
- Grupları 0 (kadın) ve 1 (erkek) olarak tanımlayın.
- Mann-Whitney U testini işaretleyin.
- Options sekmesinden Descriptives'i işaretleyin.
- OK'e tıklayın.

Mann-Whitney Testi Sonucu

Ranks

cinsiyet	N	Mean Rank	Sum of Ranks
yazma puan erkek	91	85,63	7792,00
kadin	109	112,92	12308,00
Total	200		

Test Statistics ^a

	yazma puani
Mann-Whitney U	3606,000
Wilcoxon W	7792,000
Z	-3,329
Asymp. Sig. (2-tailed)	,001

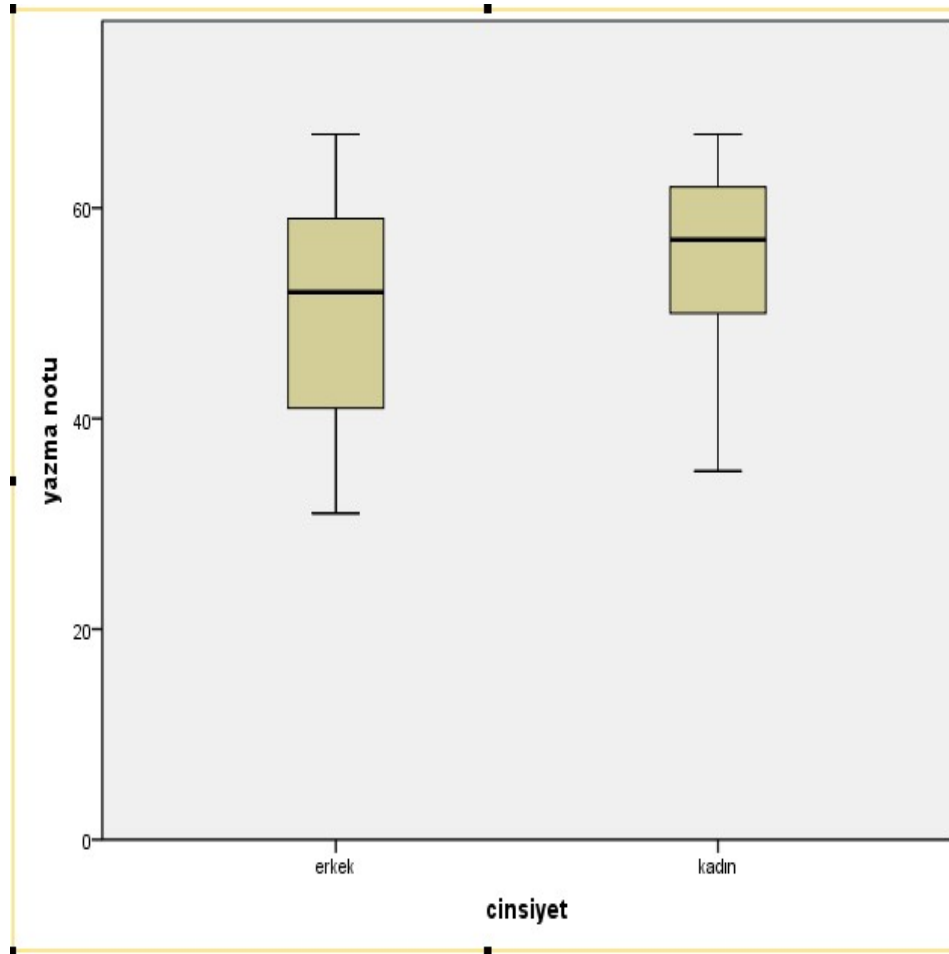
a. Grouping Variable: cinsiyet

Tabloların Yorumu

- Mann-Whitney U testinde bütün öğrencilerin yazma puanları en düşük puandan en yüksek puana doğru sıralanır.
- Erkek ve kız öğrencilere ait puanların sıraları ayrı ayrı toplanıp ortalaması alınır.
- Erkeklerin aldığı notların sıralama ortalaması 85,63, kızlarınkı 112,92.
- Yani kızlar daha yüksek puan almışlar -ki puan sıralamalarının ortalaması erkeklerinkinden daha yüksek ve aradaki fark istatistiksel açıdan anlamlı ($z = -3,329$, $p = 0,001$).
- Mann-Whitney U testi z dağılımını (SND) kullanıyor. Örneklem büyüklüğü arttıkça sonuçlar daha güvenilir olur.
- z değeri (-3,329) erkeklerin aldıkları puanların sıralarının ortalamasının 3 standart sapma altında olduğunu göstermektedir.
- Yani erkeklerin puanı istatistiksel açıdan anlamlı derecede kızlarınkinden farklıdır.
- Kız ve erkeklerin notları arasında gerçekte fark olmayıp da erkeklerin bu puanı şans eseri alma olasılıkları binde birden az (yani çok düşük).

Böylece boş hipotez reddedilir.

Boxplot



- Grafik karşılaştırma yapmak amacıyla erkek ve kızların sıralanmış notlarının ortancalarını vermektedir.
- Mann-Whitney U testi ortalamalar arasındaki farkı değil de sıra farkını ölçtüğü için grafikteki kutucuğun ortasındaki değerler ortancadır.
- Erkek ve kızların notlarının ortancaları ile %95 güven aralıkları birbirinden farklıdır.

Etki Büyüklüğü

- z değeri için etki büyüklüğü $r = Z / \sqrt{N}$ formülüyle bulunur ($r = -3,329 / \sqrt{200} = -0,23$)
- $r = -0,23$ etki büyüklüğü erkeklerle kızların notları arasında çok büyük fark olmadığını göstermektedir.

Mann-Whitney Testi Sonucunu Rapor Etme

Mann-Whitney U testi sonucu APA stiline göre şöyle rapor edilir:

- “Mann-Whitney U testi sonucuna göre erkeklerle kızların yazma notları arasında anlamlı bir fark vardır ($U = 3606,000$ $p=0,001$, $z=-3,329$, $r=-0,23$).”

Neden Daha Önce Parametrik Olmayan Test Kullanılmadı?-I

- Bu veriler için hem parametrik hem de parametrik olmayan testlerde erkek ve kızların yazma puanları arasındaki fark anlamlı çıkmıştı:
 - Mann-Whitney U test sonucu: $U = 3606$, $p=0,001$, $z=-3,329$, $r=-0,23$.
 - Bağımsız örneklem t-testi sonucu: $t = -3,66$, $SD = 169,7$, $p = 0,000$, $r=0,27$.
- Peki, yazma puanları normal dağılmamasına karşın neden aynı veriler için daha önce parametrik test kullanıldı? Çünkü:
 - Parametrik testler daha güçlüdür.
 - Parametrik testlerde **Tür 2 hatası** (gerçekte gruplar arasında fark varken fark olmadığını kabul etme olasılığı) yapma olasılığı daha düşüktür.

Neden Daha Önce Parametrik Olmayan Test Kullanılmadı?-II

- İki testin p değerlerini karşılaştırın (M-W-U testi için $p=0,001$, t -testi için $p=0,000$) .
- Bu veriler için farketmedi, iki test de iki grup arasındaki farkı yakaladı.
- Ama sınır değere (0,05) yakın başka testlerde parametrik test sonucunda iki grup arasındaki puan farkı anlamlı, parametrik olmayanda anlamsız çıkabilir ve ilkinde boş hipotez reddedilirken, ikincisinde kabul edilebilirdi.
- **Parametrik testlerin iki grup arasındaki farkı farketme gücü parametrik olmayanlarıkinden daha yüksektir.**
- **Bu nedenle veriler normal dağılıma uygun ve aralıklı oranlı ölçüm düzeyinde toplandıysa parametrik testler kullanılmalıdır.**

Wilcoxon İşaretli Sıralar Toplamı Testi

- Bağımlı (eşli) örneklem t - testinin **parametrik olmayan** karşılığıdır.
- İki değişkene ait **verilerin normal dağılması** gerekmez.
- Veriler sıralama ölçme düzeyinde toplanmış olmalı ya da aralıklı/oranlı veriler sıralama verisine çevrilmelidir (çevirme işlemi test seçildiğinde PASW yapıyor).
- Okuma ve yazma puanlarının normal dağılmadığı daha önce K-S testiyle test edildiğine göre, aynı örnek kullanılarak öğrencilerin okuma ve yazma puanları arasında fark olup olmadığı da test edilebilir.
- Araştırma hipotezi (H_1): “Öğrencilerin okuma ve yazma puanları birbirinden farklıdır.” (çift kuyruk testi).

Wilcoxon-İşaretli Sıralar Toplamı - PASW

Menüden:

- Analyze -> Nonparametric Tests-> Legacy Dialogs ->2 Related Samples'ı seçin.
- Test değişken çiftine okuma ve yazma puanlarını seçin.
- Test türüne Wilcoxon'u işaretleyin.
- Options seçeneğine tıklayarak "Descriptives"teki işareti kaldırın.
- OK'e tıklayın.

Wilcoxon İşaretli Sıra Toplamı Sonucu

Wilcoxon Signed Ranks Test

		Ranks		
		N	Mean Rank	Sum of Ranks
yazma notu - okuma notu	Negative Ranks	88 ^a	90,27	7944,00
	Positive Ranks	97 ^b	95,47	9261,00
	Ties	15 ^c		
	Total	200		

a. yazma notu < okuma notu
b. yazma notu > okuma notu
c. yazma notu = okuma notu

Test Statistics^b

	yazma notu - okuma notu
Z	-,903 ^a
Asymp. Sig. (2-tailed)	,366

- a. Based on negative ranks.
b. Wilcoxon Signed Ranks Test

Tabloların Yorumu

- Bağımlı örneklem t- testinde olduğu gibi Wilcoxon işaretli sıralar toplamı testi de öğrencilerin okuma ve yazma puanları arasında istatistiksel açıdan anlamlı bir fark olmadığını göstermektedir (Okuma puanı sıraları ort = 95,47; Yazma puanı sıraları ort= 90,27; $z = -0,903$, $p = 0,366$).

Böylece boş hipotez kabul edilir.

- Bu test öğrencilerin okuma ve yazma puanlarının ortalamaları karşılaştırılarak yapılmaz. Bütün öğrencilerin okuma ve yazma puanları sıralanır. Bir puanın diğerinden küçük, büyük ve diğerine eşit olduğu vaka sayıları saptanır.
- Örnekte yazma puanı sırası okuma puanı sırasından düşük olan 88 öğrenci (negative ranks), yazma notu sırası okuma puanı sırasından büyük olan (positive ranks) 97 öğrenci ve yazma ve okuma notu sıraları eşit olan (ties) 15 öğrenci var.

Tabloların Yorumu-2

- Negatif ve pozitif sıraların toplam ve ortalamaları veriliyor.
- Z değeri bize öğrencilerin aldıkları okuma puanlarının sıralarının ortalamasının yazma puanlarının sıralarının ortalamasından yaklaşık 5 puan düşük olduğunu gösteriyor. Bu, anlamlı bir fark değil.
- Wilcoxon işaretli sıralar toplamı testi katsayısı (T) olarak için negatif sıraların toplamı alınıyor.
- Etki büyüklüğü ($r = z / \sqrt{N} = 0,06$).

Wilcoxon İşaretli Sıralar Toplamı Testi Sonucunu Rapor Etme

Wilcoxon işaretli sıralar toplamı testi T ile gösterilir ve APA stiline göre şöyle rapor edilir:

- “Wilcoxon işaretli sıralar toplamı testi sonucuna göre öğrencilerin okuma ve yazma puanları arasında istatistiksel açıdan anlamlı bir fark yoktur ($T = 7944$, $p=0,366$, $z = -0,903$, $r=-0,06$).”
- Bağımlı örneklem t-testinde de aynı sonucu elde etmiştik (t-testindeki ortalamalar arasındaki yarım puanlık fark, bu testteki sıralar ortalaması arasındaki yaklaşık 5 puanlık fark anlamlı değil).

Wilcoxon İşaretli Sıra Testi

- İki deęişken arasındaki sıralı olarak deęil de negatif ve pozitif olarak sınıflanmıřsa Wilcoxon işaretli sıra testi yerine Wilcoxon işaret testi yapılabilir. İşaret testinde farkın sıralı olmadığı varsayılır.
- Örnek olarak okuma ve yazma verileri üzerinde işaretli sıra testi yapalım:

Wilcoxon İşaret Testi

Sign Test

Frequencies		N
yazma notu - okuma notu	Negative Differences ^a	88
	Positive Differences ^b	97
	Ties ^c	15
	Total	200

a. yazma notu < okuma notu
b. yazma notu > okuma notu
c. yazma notu = okuma notu

Test Statistics^a

	yazma notu - okuma notu
Z	-,588
Asymp. Sig. (2-tailed)	,556

a. Sign Test

- İşaret testi de puanlar arasında anlamlı fark olmadığını gösteriyor ($z = -0,588$, $p = 0,556$, $r = 0,04$).

Boş hipotez kabul edilir.

Kruskal-Wallis Testi

- Tek yönlü ANOVA testinin parametrik olmayan karşılığıdır.
- Veriler normal dağılmamışsa, grup/koşul sayısı üç ve daha fazlaysa ve tüm gruplarda/koşullarda farklı denekler kullanıldıysa **Kruskal-Wallis testi uygulanır**.
- Örneğin «öğrencilerin yazma puanları mezun oldukları lise türüne göre farklı mıdır?», sorusuna yanıt arayalım.
- Araştırma hipotezi (H_1) : “Öğrencilerin yazma puanları mezun oldukları lise türüne göre farklıdır.” $H_1: \mu \neq \mu_0$ (çift kuyruk testi)

Kruskal-Wallis Testi - PASW

Mönüden:

- Analyze -> Nonparametric Tests-> Legacy Dialogs ->K Independent Samples'ı seçin.
- Test değişkenine yazma puanını atayın.
- Gruplama değişkenine program türünü atayın ve Define range'e tıklayarak minimum 1, maksimum 3 değerlerini girin.
- Test türü için Kruskal Wallis'i ve Median'ı işaretleyin.
- Options'a tıklayarak Descriptives'i işaretleyin.
- OK'e tıklayın.

Kruskal-Wallis Testi Sonucu

Kruskal-Wallis Test

Ranks

program türü		N	Mean Rank
yazma notu	genel	45	90,64
	anadolu	105	121,56
	meslek	50	65,14
	Total	200	

Test Statistics^{a,b}

	yazma notu
Chi-square	34,045
df	2
Asymp. Sig.	,000

a. Kruskal Wallis Test

b. Grouping Variable: program türü

Median Test

Frequencies

		program türü		
		genel	anadolu	meslek
yazma notu	> Median	17	65	11
	<= Median	28	40	39

Test Statistics^b

	yazma notu
N	200
Median	54,00
Chi-square	23,456 ^a
df	2
Asymp. Sig.	,000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 20,9.

b. Grouping Variable: program türü

Tabloların Yorumu-1

- Mann-Whitney U testindeki gibi puanlar sıralanır.
- Lise türüne göre sıraların ortalaması alınır.
- K-W test istatistiği H olarak bilinir ama PASW'de ki-kare olarak veriliyor.
- Öğrencilerin mezun oldukları lise türüne göre yazma notları arasındaki fark anlamlı ($H = 34,045$, $SD=2$, $p=0,000$)

Tabloların Yorumu-2

- Ortanca (median) testi de aynı sonucu veriyor.
- Anadolu liselerinden mezun olan öğrencilerin çoğu ortancadan (54) daha yüksek not almışlar.
- Parametrik olmayan testler için yaygın kullanılan post hoc testleri yok.
- Farklı grubu/grupları bulmak için 3 Mann-Whitney testi yapmak gerekir.

Mann-Whitney Post hoc Testleri

anadolu - genel

Test Statistics^a

	yazma notu
Mann-Whitney U	1607,000
Wilcoxon W	2642,000
Z	-3,111
Asymp. Sig. (2-tailed)	,002

a. Grouping Variable: program türü

anadolu - meslek

Test Statistics^a

	yazma notu
Mann-Whitney U	1169,000
Wilcoxon W	2444,000
Z	-5,587
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: program türü

meslek- genel

Test Statistics^a

	yazma notu
Mann-Whitney U	813,000
Wilcoxon W	2088,000
Z	-2,331
Asymp. Sig. (2-tailed)	,020

a. Grouping Variable: program türü

- Üç farklı grup için test yapılır; her birinin kendi alfa yanılma yüzdesi (α) olduğundan 0,05 üçe bölünür (0,0167) (buna Bonferroni düzeltmesi denilir).
- Yani meslek lisesiyle genel lise puanları arasındaki fark anlamlı olmaktan çıkar (0,02 > 0,017).
- ANOVA testinde de ikisi arasındaki fark sınırın biraz üstündeydi.

Kruskal-Wallis Testi Etki Büyüklüğü

- Kruskal Wallis testi için etki büyüklüğü her ikili grup için ayrı ayrı hesaplanır.
- Anadolu liseleri ile meslek liseleri için $r = -5,587 / \sqrt{(105+50)} = -0,45$ (yani 0,5'e yakın olduğu için etki katsayısı büyük sayılır).
- $r_{\text{anadolu-genel}} = -0,25$ (etki küçük)
- $r_{\text{meslek-genel}} = -0,24$ (etki küçük)

Kaynak: Field ve Hole, 2008, s. 247-249

Kruskal Wallis Testi Sonucunu Rapor Etme

Kruskal Wallis testi APA stiline göre şöyle rapor edilir:

- “Öğrencilerin yazma notları hangi tür liseden mezun olduklarına göre anlamlı düzeyde farklılık göstermektedir ($H = 34,045$, $SD=2$, $p=0,000$). Farkın hangi grup ya da gruplardan kaynaklandığını bulmak için Mann-Whitney U testleri yapıldı. Bonferroni düzeltmesi uygulanarak tüm etkiler için anlamlılık düzeyi $0,0167$ olarak kabul edildi. Meslek liseleriyle genel lise mezunlarının yazma puanları arasında anlamlı bir fark gözlenmedi ($U=813$, $p > 0,0167$). Anadolu liseleriyle meslek liselerinden ve anadolu liseleriyle genel liselerden mezun olan deneklerin yazma notlarının birbirinden anlamlı düzeyde farklı olduğu görüldü (sırasıyla $U=1169$, $r=-0,45$, $U=1607$, $r=-0,25$). Anadolu liselerinden mezun olan öğrencilerin yazma puanları hem meslek liselerinden hem de genel liselerden mezun olanlardan daha yüksektir.”

Boş hipotez reddedilir.

Kaynak: Field ve Hole, 2008, s. 247-249



Friedman'ın ANOVA Testi

Tek yönlü **tekrarlı** ANOVA testinin parametrik olmayan karşılığıdır.

- Veriler normal dağılmamışsa, **grup/koşul sayısı üç ve daha fazlaysa** ve **tüm gruplarda/koşullarda aynı denekler** kullanıldıysa Friedman'ın ANOVA testi uygulanır.
- Örneğin «öğrencilerin okuma, yazma ve matematik puanları birbirinden farklı mıdır?», sorusuna yanıt arayalım.
 - Araştırma hipotezi (H_1) : “Öğrencilerin okuma, yazma ve matematik puanları birbirinden farklıdır.” (çift kuyruk testi)
 - Boş hipotez (H_0) her dersin puan türü sıralarının benzer olup olmadığını test eder

Önce Kolmogorov-Smirnov Normallik Testi

Mönüden:

- Analyze -> Descriptive Statistics-> Explore'u seçin.
- Bağımlı değişken listesine okuma, yazma ve matematik puanlarını atayın.
- Both seçeneği işaretli (İstatistikler ve Grafik için).
- Statistics sekmesinde Descriptives'in ve %95 güven aralığı seçili (değiştirilebilir).
- Plots (grafik) Normality plots with test seçeneğini işaretleyin
- OK'e tıklayın.

K-S Normallik Testi Sonucu

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
yazma notu	,134	200	,000	,947	200	,000
okuma notu	,105	200	,000	,980	200	,006
matematik notu	,071	200	,016	,977	200	,002

a. Lilliefors Significance Correction

- K-S testi sonuçları yazma, okuma ve matematik puan türleri için de anlamlı (sırasıyla $D(200)=0,134$, $p=0,000$, $D(200)=0,105$, $p=0,000$), $D(200)=0,071$, $p=0,016$).
- Yani her üç puan türü için de verilerin dağılımı normal değil, dolayısıyla **parametrik olmayan test uygulanabilir**.

Friedman'ın ANOVA Testi - PASW

Menüden:

- Analyze -> Nonparametric Tests-> Legacy Dialogs ->K Related Samples'ı seçin.
- Test değişkenlerine okuma, yazma ve matematik puanlarını atayın.
- Test türü için Friedman'ı işaretleyin.
- Statistics'e tıklayarak Descriptive ve Quartile'ı işaretleyin.
- OK'e tıklayın.

Friedman'ın ANOVA Testi Sonucu

Descriptive Statistics								
	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
yazma notu	200	52,78	9,479	31	67	45,25	54,00	60,00
okuma notu	200	52,23	10,253	28	76	44,00	50,00	60,00
matematik notu	200	52,65	9,368	33	75	45,00	52,00	59,00

Friedman Test

Ranks

	Mean Rank
yazma notu	2,04
okuma notu	1,96
matematik notu	2,01

Test Statistics^a

N	200
Chi-square	,645
df	2
Asymp. Sig.	,724

a. Friedman Test

Tabloların Yorumu

- İlk tabloda tanımlayıcı istatistikler verilmiş. Parametrik olmayan test yapıldığı için **ortanca** (median) önemli.
- Test istatistiği ki- kare ile verilmiş, SD 2, p değeri 0,724 ($\chi^2_{(2)} = 0,645, p = 0,724$).
- Test sonucuna göre üç puan türünün dağılımları birbirine benziyor. Sıra ortalamaları ve ortancalar birbirine çok yakın.
- **Boş hipotez kabul edilir:** Öğrencilerin okuma, yazma ve matematik puanları birbirine benzemektedir.
- Boş hipotez reddedilmiş olsaydı daha önce Kruskal Wallis testinde olduğu gibi post hoc testleri (Wilcoxon) yapmak ve etki büyüklüğünü hesaplamak gerekecekti. (Parametrik olmayan testler için yaygın kullanılan post hoc testleri olmadığı daha evvel ifade edilmişti).

Friedman'ın ANOVA Testi Sonucunu Rapor Etme

Friedman'ın ANOVA testi sonucu APA stiline göre şöyle rapor edilir:

- “Friedman'ın ki- kare testi sonucu öğrencilerin okuma, yazma ve matematik notları arasında anlamlı bir fark olmadığını göstermektedir ($\chi^2_{(2)} = 0,645, p = 0,724$). Her üç derse ait puanların dağılımı birbirlerine benzemektedir.”

Boş hipotez kabul edilir.

Parametrik Olmayan Korelasyon Testi (Spearman's ρ)

- İki değişken arasındaki ilişkinin gücünü ölçmek için kullanılan parametrik olmayan bir testtir.
- Verilerin normal dağılmış olması gerekmez.
- Sıralama ölçüm düzeyinde toplanmış veriler için kullanılır.
- Veriler aralıklı ya da oranlı ölçüm düzeyinde toplanmışsa önce sıralamaya çevrilir, sonra test uygulanır.
 - Yine okuma ve yazma puanları arasındaki korelasyona bakalım. Verilerin normal dağılmadığı bilinmektedir:
 - Araştırma hipotezi (H_1): “Öğrencilerin okuma ve yazma puanları arasında bir korelasyon vardır.” ($H_1 : \mu_1 \neq \mu_2$) (çift kuyruk testi).

Korelasyon Testi - PASW

Menüden:

- Analyze -> Correlate-> Bivariate'i seçin.
- Yazma ve okuma puanlarını atayın.
- Spearman ve two-tailed test'i seçin.
- OK'e tıklayın.

Spearman Korelasyon Testi Sonucu

Correlations

			okuma notu	yazma notu
Spearman's rho	okuma notu	Correlation Coefficient	1,000	,617**
		Sig. (2-tailed)	.	,000
		N	200	200
	yazma notu	Correlation Coefficient	,617**	1,000
		Sig. (2-tailed)	,000	.
		N	200	200

** . Correlation is significant at the 0.01 level (2-tailed).

- Veriler sıralı olduğu için Spearman korelasyon katsayısında aritmetik ortalama ya da standart sapma seçenekleri hesaplanmaz.

Tablonun Yorumu

- Öğrencilerin okuma ve yazma puanları arasında pozitif bir korelasyon (0,617) var ve bu korelasyon anlamlıdır (0, 000).
- Parametrik olmayan Spearman korelasyon katsayısı Spearman's ρ ya da Spearman's ρ ile gösterilir.
- Korelasyon katsayısı okuma ve yazma puanları arasındaki korelasyonun yaklaşık 0,38'ini açıklıyor (ρ 'nun karesi alınır).
- Yani okuma puanlarının %38'i yazma puanlarındaki değişimle açıklanabilir.
- Yani okuma puanları yüksek olan öğrencilerin yazma puanları da yüksektir (ya da yazma puanları yüksek olan öğrencilerin okuma puanları da yüksektir).

Boş hipotez reddedilir.

Spearman Korelasyon Testi Sonucunu Rapor Etme

APA stiline göre bulgular şöyle rapor edilir:

- “Öğrencilerin okuma ve yazma puanları arasında pozitif bir korelasyon gözlenmiştir (Spearman’s $\rho = 0,617$, $p = 0,000$, $\rho^2 = 0,38$). İki değişken arasında orta düzeyde güçlü bir korelasyon vardır. Okuma puanları yüksek olan öğrencilerin yazma puanları da nispeten daha yüksektir.”
- Son cümle “Yazma puanları yüksek olan öğrencilerin okuma puanları da nispeten daha yüksektir.” şeklinde de yazılabilir.

Parametrik ve Parametrik Olmayan Korelasyon Testi Karşılaştırması

Correlations

		okuma notu	yazma notu
okuma notu	Pearson Correlation	1	,597**
	Sig. (2-tailed)		,000
	N	200	200
yazma notu	Pearson Correlation	,597**	1
	Sig. (2-tailed)	,000	
	N	200	200

** . Correlation is significant at the 0.01 level (2-tailed).

Correlations

			okuma notu	yazma notu
Spearman's rho	okuma notu	Correlation Coefficient	1,000	,617**
		Sig. (2-tailed)	.	,000
		N	200	200
	yazma notu	Correlation Coefficient	,617**	1,000
		Sig. (2-tailed)	,000	.
		N	200	200

** . Correlation is significant at the 0.01 level (2-tailed).

- Pearson korelasyon testi ile Spearman aynı sonucu verdi (yani iki değişken arasında korelasyon var).
- Ama Spearman korelasyon katsayısı Pearson'dan daha yüksek.
- Çünkü parametrik olmayan testler parametrik testlerden daha az duyarlıdır.

Özet

- Parametrik olmayan testler
 - Binom testi
 - Ki-kare testi
 - Mann-Whitney testi
 - Wilcoxon işaretli sıralar toplamı ve işaret testi
 - Kruskal Wallis testi
 - Friedman'ın ANOVA testi
 - Parametrik olmayan korelasyon testi

Bu ders sunumu, Prof. Dr. Yaşar TONTA'nın 29.09.2021 tarihli izniyle tarafımızdan düzenlenerek kullanılmıştır. Kendisine sunumun ana çatkısı ve içeriği için teşekkür ederiz.