**T.C.**

|  |
| --- |
|  |

**GAZİ ÜNİVERSİTESİ**

**EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**ÖLÇME DEĞERLENDİRME BÖLÜMÜ**

**YÜKSEK LİSANS PROGRAMI**

**2022-2023 EĞİTİM ÖĞRETİM YILI GÜZ DÖNEMİ**

**R İLE İSTATİSTİKSEL PROGRAMLAMA DERSİ**

**VİZE ÖDEVİ**

**1-0 PUANLANMIŞ MADDELERİN R UYGULAMALARI**

**Hazırlayan:**

22810414523

USLU ÇELİK GÜRLEYEN

# İÇİNDEKİLER

[İÇİNDEKİLER 1](#_Toc124340503)

[BÖLÜM 1: R’A GİRİŞ 1](#_Toc124340504)

[**1.** **1. R Genel Bakış** 1](#_Toc124340505)

[**1.2.R Kaynakları ve R’ın Yüklenmesi** 2](#_Toc124340506)

[**1.3.Neden R kullanıyoruz?** 4](#_Toc124340507)

[BÖLÜM 2. ÖRNEK R UYGULAMASI 5](#_Toc124340508)

[**2.1. Verinin Okutulması** 5](#_Toc124340509)

[**2.2 Madde Analizleri** 9](#_Toc124340510)

[**2.3. Madde Güçlük Düzeyi** 9](#_Toc124340511)

[**2.4. Madde Ayırt Edicilik Düzeyi** 11](#_Toc124340512)

[**2.5. Madde Güvenirlik Analizi** 14](#_Toc124340513)

[**2.6. OUT Oluşturma** 15](#_Toc124340514)

[BÖLÜM 3. VERİ GÖRSELLEŞTİRMELERİ 16](#_Toc124340515)

[**Grafiklerin Kaydedilmesi** 22](#_Toc124340516)

[KAYNAKÇA 26](#_Toc124340517)

# BÖLÜM 1: R’A GİRİŞ

## **1. R Genel Bakış**

R, 1970lerin sonunda Bell laboratuvarlarında geliştirilen S programının devamı niteliğinde, istatistiksel hesaplamaların yapıldığı bir programlama dilidir. İsmi programın geliştiricileri Ross Ihaka ve Robert Gentleman’ın baş harflerinden gelmektedir. John Chambers ve arkadaşları tarafından geliştirilen “S” dilinin açık kaynak kodlu halidir. R programı, dinamik bir yazılımdır. Açık kaynak kodunun geliştirilmesinde büyük bir topluluğa sahiptir. 1997 yılında aralarında Ihaka, Gentleman, Mächler, Peter Dalgaard, Kurt Hornik, Friedrich Leisch ve Thomas Lumley dahil 11 kişi R’nin kaynak kodunun anahtarlarına erişerek kendilerini “R Core Team” olarak adlandırdı. Bu ekibe dair daha fazla bilgiye erişmek isterseniz R programının sitesinden (<https://www.r-project.org/>) ulaşabilirsiniz.

Ücretsiz açık kod sistemi ile çalışan veri analizi ve veri madenciliğinde kullanılan bir istatistik programıdır. 1990lı yıllarda geliştirilmeye başlanmış 2000 yıllarından itibaren geniş kitlelerin erişimine açılmıştır. Çok geniş skalada kullanım alanı ve ücretsiz oluşu R’ın tercih edilme sebeplerini artırmaktadır. Çok geniş alandan hemen her sektörde kullanımında anlamlı analizler yapılabilmesi için içerisinde sürekli güncellenen versiyonları ile birçok teknik paket mevcuttur. İhtiyaç duyulan analiz yöntemlerine farklı paketlerdeki farklı işlemler ile cevap bulunabilmektedir.

Temel düzeyde R base paketi ile temel istatistikler hesaplanabilmektedir.

## **1.2.R Kaynakları ve R’ın Yüklenmesi**

R programında çalışabilmek için öncelikle programın indirilmesi gerekmektedir. Programın farklı web adreslerinden erişimi mümkündür. Tavsiye edilen linkler kullanılarak indirme işlemi gerçekleştirilebilir. Güncel sürümler ve bilgisayar özelliklerine dikkat edilerek indirme işlemi yapılmasına dikkat edilmelidir.

**Official R web page :** [**http://cran.r-project.org**](http://cran.r-project.org)[**/**](http://cran.r-project.org/)

**First Cran.r then Rstudio :** [**http://www.rstudio.com/ide/download/**](http://www.rstudio.com/ide/download/)

R’nin yüklenmesi için cran.r-project.org sayfasına girildiğinde “Download and Install R” yazılı bir ekran çıkacaktır. Burada işletim sistemine göre yüklenecek tür seçildikten sonra gelen ekranda, ilk defa yüklenecekse base yazısına veya install R for the first time yazısına tıklanarak gelen ekrandan R’ın en son versiyonu yüklenebilir. RStudio’nun yüklenmesi için de RStudio.com adresine girilerek Products kısmından RStudio seçilir ve ücretsiz olan RStudio Desktop FREE seçilerek yükleme yapılır. RStudio’da dört temel bölüm bulunmaktadır. Bunların konumları kullanıcılar tarafından ayarlanabilmekle birlikte, kullanıcıya göre sol üst tarafta kod yazım editörünün bulunduğu Source, sağ üst tarafta kodların değerlendirilmesini ve sonuçları gösteren Console, sol alt tarafta veri setlerinin ve değişkenlerin bulunduğu Environment ve sağ alt tarafta grafikler, paketler ve yardım dosyalarına erişim imkanının olduğu Help bölümleri bulunmaktadır.

Sosyal bilimlerde yaygın olarak kullanılan paketlerden yaptığımız işlemler için base temelinde gerekli paketler:

* **R Programlama :** installr
* **Grafik:** ggplot2, graphics, lattice, plotly
* **Veri düzenleme:** data.table, tidyr, dplyr, tidyverse
* **Veri yükleme:** readr, readxl, XML

İstenilen paketlerin yüklenebilmesi için konsol kısmında install.packages() fonksiyonunda parantez içerisine paket adı yazılmalıdır.

install.packages(readr)

Paketler yüklendikten sonra library() fonksiyonu kullanılmalıdır. Ayrıca R Studıo ekranında sağ tarafta bulunan bölümde Packages kısmından işaretleme yapılarak da paket işler hale getirilebilir.

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

Paket yüklemesi bir kez yapılırken library() fonksiyonu her çalışmada yenilenmelidir. Aksi durumda hata mesajı alınmaktadır.

## **1.3.Neden R kullanıyoruz?**

R programlama dili, bilimsel araştırmalarda verileri temizleyip, analiz edip, görselleştirerek anlamlı hale getirebildiğimiz ücretsiz bir yazılımdır. Oysa istatiksel çalışmalarda sıklıkla tercih edilen SPSS, SAS gibi programlar ücretlidir ve birtakım psikometri analizleri için yeterli değildir. MPlus, IRTPRO, HLM gibi özel psikometrik analizler için geliştirilen programlar olsa dahi bu programlar oldukça maliyetli olduğundan R diğer programlara ciddi bir rakiptir.

R programı kişiselleştirilip geliştirebileceğiniz zengin bir grafik imkanına sahiptir.

Oluşturulan bu araç kutusu(toolbox), istatistiksel analizler yaparak verilerini görselleştirmek isteyen veri bilimcilere ve ölçme ve değerlendirme alanında çalışmalar yapan kişilere hizmet edecek şekilde düzenlenmiştir.

# BÖLÜM 2. ÖRNEK R UYGULAMASI

## **2.1. Verinin Okutulması**

Bu bölümde KTK analizlerinin R’da gerçekleştirilmesi, çalışma dizininde kayıtlı “DATA\_15.xlsx” olarak adlandırılan veri seti üzerinden örneklerle anlatılmıştır. Bu veri setinin görüntüsü Şekil 1’de yer almaktadır.

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

*Şekil 1. “DATA\_15” Verisinin Görüntüsü*

“DATA\_15.xlsx” veri seti 15 çoktan seçmeli maddeden oluşan bir testi alan 1000 öğrencinin yanıtlarını içermektedir.

“DATA\_15.xlsx” veri dosyası R’da *read\_excel ()* fonksiyonu kullanılarak aşağıdaki komutla okunabilir. Dosyanın okutulması için önce *library (readxl)* çağrılmalıdır. Komutun çıktısı olarak “Data” adıyla kaydedilen nesne 1000x15 boyutlarında bir data framedir.

|  |
| --- |
| library(readxl)  Data <- read\_excel("C:\\AR\\RVİZEDATA\\DATA\_15.xlsx", col\_names = FALSE ) |

“Data” nesnesi matrise dönüştürme işlemi için *as.matrix()* komutu kullanılmıştır.

|  |
| --- |
| #matrise dönüştürme  Data<-as.matrix(Data) |

“Data” matrisinin sütun sayısı ve satır sayısı sırasıyla testteki madde sayısını (*k*) ve testi alan birey sayısını(*n*) vermektedir.

|  |
| --- |
| # n ve k değeri tanımlama  k<- ncol(Data)  k  n <- nrow(Data)  n |

“Data” nesnesinin satır toplamları alınarak öğrencilerin toplam puanları elde edilebilmektedir. Aşağıdaki komutlarda ilk olarak “Data” nesnesinin satır toplamlarını hesaplamk için *rowSums*() fonksiyonu kullanılarak bireylerin puanları “X” adıyla kaydedilmiştir. Sonra *mean()* fonksiyonu kullanılarak toplam puanların ortalaması hesaplanmış ve toplam puanların ortalaması “Xort” adıyla kaydedilmiştir. Daha sonra varyans formülü kullanılarak toplam puanların varyansı, varyansın karekökü hesaplamaları yapılarak toplam puanların standart sapması hesaplanmıştır ve toplam puanların varyansı “X.VAR”, standart sapması ise “X.STDSPM” adıyla kaydedilmiştir.

|  |
| --- |
| tablo içeren bir resim  Açıklama otomatik olarak oluşturuldu    #satır toplamları  X<- rowSums (Data)  X  #data dosyasına toplam puan ekleme  Data<- cbind (Data , X)  #toplam puan ortalaması  Xort<- mean(X)  Xort  ## [1] 6.297  #toplam puan varyansı  X.VAR<- sum((X-X.Xort)^2)/ (length(X))  X.VAR  ##[1] 7.906791  #toplam puan st.sapma  X.STDSPM <- sqrt(X.VAR)  X.STDSPM  ##[1] 2.811902 |

Buna göre toplam puanların ortalaması 6.297, standart sapması ise 2.811 olarak hesaplanmıştır.

## **2.2 Madde Analizleri**

Bireylere ilişkin madde puanlarının yer aldığı “Data” nesnesi üzerinden madde analizleri ve güvenirlik analizleri geçekleştirirken izlenilen adımlar ve kullanılacak komutlar aşağıdaki gibidir.

Madde analizleri bireylerin bir testte yer alan her bir maddeye ilişkin yanıtlarının istatistiksel özelliklerinin hesaplanmasını ve incelenmesini tanımlamak için kullanılan bir terimdir (Crocker & Algina , 1986).

Madde puanlarının dağılımını tanımlayan ortalama ve standart sapma gibi betimsel istatistikler ile testteki hangi maddelerin daha kullanışlı olabileceği ile ilgili yorumlar yapılabilir. Ortalaması dağılımın merkezine daha yakın olan ve değişkenliği fazla olan maddelerin genellikle daha iyi performans göstermesi beklenmektedir.

## **2.3. Madde Güçlük Düzeyi**

İki kategorili puanlanan maddelerin madde puanları ortalaması maddeyi dğru yanıtlayan bireylerin oranıdır. Bu orana madde güçlük düzeyi denir. j maddesinin madde güçlük düzeyi pj olarak gösterilir. “Data” nesnesindeki madde puanlarının ortalaması *colMeans()* fonksiyonu kullanılarak aşağıdaki komut ile hesaplanmaktadır. Komutun çıktısı “p” adı ile kaydedilen nesne 15 elemanlı bir vektördür.

|  |
| --- |
| #madde puanları ort.  p<- colMeans(Data)  p  ## 1 2 3 4 5 6 7 8 9 10 11 12  0.571 0.479 0.447 0.435 0.420 0.389 0.365 0.347 0.295 0.235 0.232 0.611  13 14 15  0.617 0.421 0.433 |

“p” nesnesinin ilk elemanı 1.maddenin güçlük düzeyini belirtmektedir ve çıktıya göre Madde 1’in güçlük düzeyi 0,571’dir. Diğer bir ifade ile grubun %57’si tarafından doğru yanıtlanmıştır. Madde güçlükleri incelendiğinde en zor maddenin 0,232 güçlük düzeyi ile 11.madde ve 0,235 güçlük düzeyi ile 10.madde olduğu görülmektedir. Testteki 13.madde grubun %61,7’si tarafından doğru cevaplanarak testin en kolay maddesidir.

İki kategorili puanlanan bir maddenin madde puanları varyansı maddeyi doğru yanıtlayan bireylerin oranının maddeyi yanlış yanıtlayan bireylerin oranı ile çarpılması işlemi ile hesap edilir.

(p\*(1-p)) ya da (p\*q)

Varyansın karekökünün alınması ile standart sapma değeri hesaplanır (√p\*q).

Aşağıdaki komutlarda ilk olarak her madde için maddeyi yanlış yanıtlayan bireylerin oranı hesaplanarak “q” nesnesi olarak tanımlanmıştır. Sonra her bir maddenin varyansı hesaplanarak “p.VAR” adıyla, daha sonra da standart sapması hesaplanarak “p.STDSPM” adı ile kaydedilmiştir.

|  |
| --- |
| #maddeyi yanls yantlynların oranı  q<- (1-p)  q  ## 1 2 3 4 5 6 7 8 9 10 11 12  0.429 0.521 0.553 0.565 0.580 0.611 0.635 0.653 0.705 0.765 0.768 0.389  13 14 15  0.383 0.579 0.567  #madde varyansları  p.VAR<- (p\*q)  p.VAR  ## ...1 ...2 ...3 ...4 ...5 ...6 ...7 ...8  0.244959 0.249559 0.247191 0.245775 0.243600 0.237679 0.231775 0.226591  ...9 ...10 ...11 ...12 ...13 ...14 ...15  0.207975 0.179775 0.178176 0.237679 0.236311 0.243759 0.245511  #madde st.spma  p.STDSPM<- sqrt(p.VAR)  p.STDSPM  ##  ...1 ...2 ...3 ...4 ...5 ...6 ...7  0.4949333 0.4995588 0.4971831 0.4957570 0.4935585 0.4875233 0.4814302  ...8 ...9 ...10 ...11 ...12 ...13 ...14 ...15  0.4760158 0.4560428 0.4239988 0.4221090 0.4875233 0.4861183 0.4937196 0.4954907 |

“p.VAR” ve “p.STDSPM” nesnelerinin elemanları her bir maddeye ait varyans ve standart sapma değerleridir. Buna göre testin 2.maddesinin varyans değeri 0.2495 , standart sapması ise 0.4995 değerleri ile testte değişkenliği en fazla olan maddedir. Testteki 11.maddenin varyansı 0.1781 , standart sapma değeri ise 0.4221 değeri ile testin değişkenliği en az olan maddesidir.

## **2.4. Madde Ayırt Edicilik Düzeyi**

Madde ayırt ediciliği madde puanları ile toplam puanlar arasındaki korelasyon katsayısının çarpımı ile hesaplanabilir. İki kategorili puanlanan bir madde için değişkenlerden birisi iki kategorili olduğunda, Pearson Momentler Çarpımı korelasyonunun bir versiyonu olan nokta çift serili korelasyon katsayısından veya çift serili korelasyon katsayısından faydalanılabilir.

Nokta çift serili korelasyon katsayısı, iki kategorili değişken için gerçekte iki kategori söz konusu olduğunda kullanılır. Bir j maddesi için nokta çift serili korelasyon katsayısı aşağıdaki formül ile hesaplanabilir:

rpbis = \*

Burada ,

µ+ : Maddeyi doğru yanıtlayan bireylerin ortalamatoplam test puanı

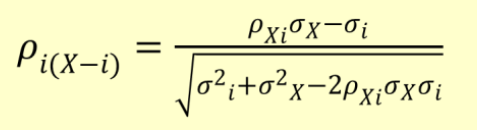
µx : Tüm grup için ortalama toplam test puanı

σx : Tüm grup için toplam test puanlarının standart sapması

pj: j maddesi için madde güçlüğü

qj : j maddesi için 1- pj

Bir maddeye ilişkin madde puanları ile toplam puanlar arasındaki nokta çift serili korelasyon katsayısının değeri madde puanlarının toplam test puanına dahil olması sebebiyle yanıltıcı olabilir. Bir testte çok sayıda madde varsa testteki herhangi bir maddenin toplam puan üzerindeki etkisi oldukça düşüktür. Ancak toplam puan üzerindeki madde etkisinin düşüklüğü için testte en az kaç madde olması gerektiğine dair bir kesinlik olmadığı için korelasyon katsayısının kestiriminde düzeltilmiş toplam puan tercih edilebilir (Klein , 2005). Bir testte 25’ten daha düşük az sayıda madde bulunduğunda düzeltilmiş nokta çift serili korelasyon katsayısı aşağıdaki gibi hesaplanabilir:



σx : Tüm grup için toplam test puanlarının standart sapması

σx 2 : Tüm grup için toplam test puanlarının varyansı

σj : Tüm grup için j maddesinin madde puanlarının standart sapması

σj 2 : Tüm grup için j maddesinin madde puanlarının varyansı

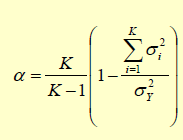
“Data” nesnesindeki madde puanları ile bireylerin toplam puanları arasındaki nokta çift serili korelasyon katsayısı ve düzeltilmiş nokta çift serili korelasyon katsayısı aşağıdaki komut ile hesaplanmaktadır. Komutta yer alan *for* döngüsünden önce “X1.ORT”, “R.pbis” ve “ “duz.R.pbis” isimleri ile boş vektörler oluşturulmuştur. *for* döngüsünde ilk olarak j maddesi için maddeyi doğru yanıtlayan öğrencilerin ortalama test puanı hesaplanmıştır. Bunun için “Data” nesnesinin j. Sütununda yer alan madde puanlarında 1 değerine sahip olan bireyler için rowSums () fonksiyonu ile satırlar toplamı alınarak toplam puanlar elde edilmiş, mean() fonksiyonu ile hesaplanan satırlar toplam puanının ortalaması hesaplanmıştır. Elde edilen toplam puan “X1.ORT” nesnesinin j. Elemanı olarak kaydedilmiştir. Ardından j maddesi için nokta çift serili korelasyon ve düzeltilmiş nokta çift serili korelasyon katsayıları hesaplanarak sırası ile “R.pbis” ve “duz.R.pbis” nesnesinin j. elemanı olarak kaydedilmiştir. Döngü her bir madde için tekrarlanmıştır. “R.pbis” ve “duz.R.pbis” isimli kayıtlı nesneler 15 elemanlı vektörler olarak kaydedilmiştir.

|  |
| --- |
| #bos vektorler olstrma  X1.ORT<- vector()  R.pbis<- vector()  duz.R.pbis<- vector()  #nokta cft serili ve düzeltilms nokta cift serili korelasyon  for( j in 1:k){  X1.ORT[j]<- mean(rowSums(Data[Data[ ,j]==1, 1:k]))  R.pbis[j] <- ((X1.ORT[j]-Xort)/ X.STDSPM)\*(sqrt(p[j]/q[j]))  duz.R.pbis[j]<- ((R.pbis[j]\*X.STDSPM)- p.STDSPM[j])/  sqrt(X.VAR+p.VAR[j]-(2\*R.pbis[j]\*X.STDSPM\*p.STDSPM[j]))  }  R.pbis  ## [1] 0.3933401 0.4212520 0.4136113 0.4123363 0.4159419 0.4175946 0.3993361  [8] 0.3197146 0.3902102 0.3272865 0.3800531 0.3687686 0.3582888 0.4026267  [15] 0.4309276  duz.R.pbis  ## [1] 0.2300408 0.2593937 0.2517131 0.2507991 0.2555904 0.2595612 0.2414629  [8] 0.1567975 0.2403974 0.1836109 0.2412485 0.2057095 0.1947904 0.2407425  [15] 0.2716542 |

“R.pbis” ve “duz.R.pbis” nesnelerinin birinci elemanları testteki 1. Madde için madde toplam puan ve madde düzeltilmiş toplam puan korelasyon katsayılarıdır. 1. Madde için madde düzeltilmiş toplam puan korelasyon katsayısı 0.230’dur. Bu indeksin değerinin artı olması, testteki belli bir maddeyi toplam puanı yüksek olanların toplam puanı düşük olanlardan daha fazla doğru cevapladığı anlamına gelir. Bu maddenin ayırt edicilik değerine göre kısmen düzeltilmesi şartı ile teste alınabileceği yorumu yapılabilir. Tüm maddelerin ayırt edicilik değerlerine bakıldığında 0.156 ayırt edicilik indeksine sahip olan 8.maddenin testten çıkarılması gerektiği söylenebilir. Diğer maddeler de incelendiğinde en yüksek ayırt edicilik değerinin 0.271 olduğu ve maddelerin çok düşük ayırt ediciliğe sahip olduğu, düzeltilmeye ihtiyaç olduğu söylenebilir.

## **2.5. Madde Güvenirlik Analizi**

Test puanlarının güvenirliği için yaygın olarak Cronbach alfa katsayı değeri hesaplanmaktadır. Cronbach alfa katsayısı hem çok kategorili puanlanmış maddelerin hem de iki kategorili puanlanmış maddelerin oluşturduğu testlerde güvenirlik yorumlanmasında kullanılabilir.



K: madde sayısı

σi 2 : Tüm grup için j maddesinin madde puanlarının varyansı

σY 2 : Tüm grup için toplam test puanlarının varyansı

“Data” nesnesindeki madde puanlarına ilişkin toplam puanların güvenirlik katsayısının R’da hesaplanması aşağıda gösterilmiştir.

|  |
| --- |
| alfa<- (k/(k-1))\* (1- ((sum(p\*q))/(X.VAR)))  alfa  ## [1] 0.6030724 |

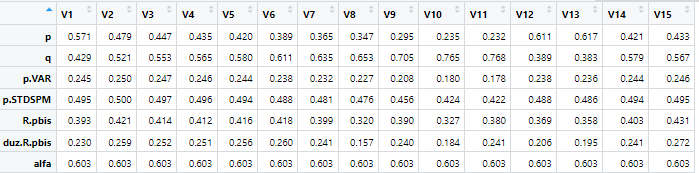
Alfa katsayısının değerine göre toplam test puanlarının güvenirliği 0.603 olarak hesaplanmıştır. 0.70’ten küçük güvenirlik değerleri düşük güvenirlik olarak yorumlanmaktadır.

## **2.6. OUT Oluşturma**

R’da gerçekleştirilen hesaplama sonuçlarını toplu bir şekilde görebilmek amacıyla out oluşturulabilir. Bunun için çalıştırılan komut ve çıktı aşağıdaki gibidir:

|  |
| --- |
| #OUT oluşturma  DataOut<- matrix(c(rep(0,n\*k)),n,k)  DataOut<- as.matrix(Data)  is.matrix(DataOut)  # Hesaplanan değerlerin yazılacağı matrisi oluşturma  DataOut<-matrix(c(rep(0,7\*15)),7,15)  colnames(DataOut) <- c(colnames(DataOut))  rownames(DataOut)<- c("p","q","p.VAR","p.STDSPM","R.pbis","duz.R.pbis","alfa")  # outyazdırma  ###outa[rownames(),col]<- round(rownames,3)  DataOut[1,]<- round(p,3)  DataOut[2,]<- round(q,3)  DataOut[3,]<- round(p.VAR,3)  DataOut[4,]<- round(p.STDSPM,3)  DataOut[5,]<- round(R.pbis,3)  DataOut[6,]<- round(duz.R.pbis,3)  DataOut[7,]<- round(alfa,3) |

R Studio ile hesaplanan KTK analizi sonuçlarının hepsini tek bir yerde görmek için oluşturulan nesnenin görüntüsü aşağıda verilmiştir.



*Şekil 2. DataOut nesnesinin görünümü*

# BÖLÜM 3. VERİ GÖRSELLEŞTİRMELERİ

Veri görselleştirmesine dair birçok paket bulunmaktadır. R temel pakette ggplot ve dplyr temel grafiklerin elde edilmesinde oldukça foksiyonel ve yeterlidir. Grafik komutu çalıştırılmadan önce “*ggplot*” ve “*dplyr*” paketlerinin yüklemesi yapılmalı ardından *library()* fonksiyonu ile paketler çalıştırılmalıdır.

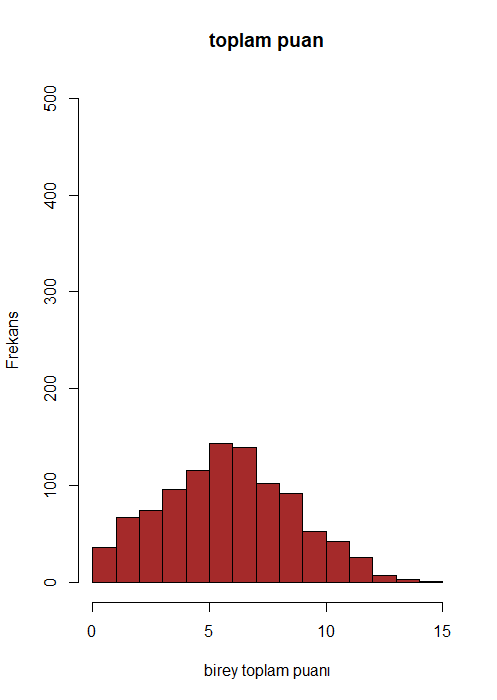
1-0 olarak puanlanmış maddelere ilişkin bireylerin toplam puanının dağılımı için histogram grafiği çizdirilebilir. Çizdirilen histogram grafiği ile bireylerin toplam puanlarının dağılımı yorumlanabilir. Histogram grafiği *hist()* fonksiyonu ile elde edilmektedir. Aşağıda bireylere ait toplam puanların histogramının komut ve çıktısı yer almaktadır.

#histogram (toplam puanlar)

hist(Data[ , 16], main = "toplam puan",

xlab = "birey toplam puanı", ylab = "Frekans", xlim = c(0,15), ylim = c(0,500),

breaks = 13, col = "brown")

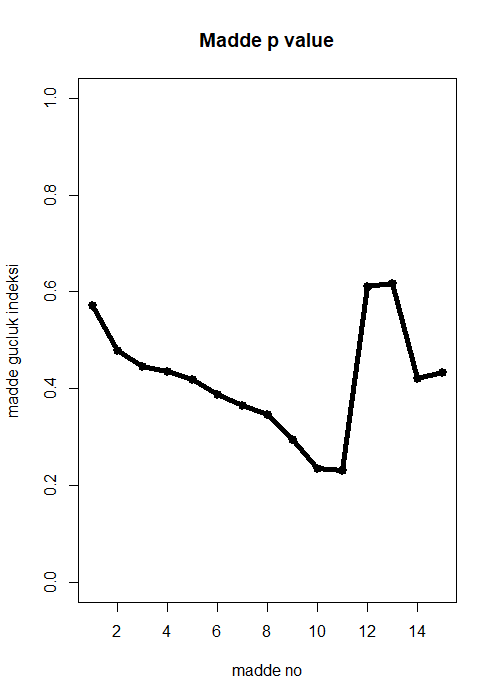


*Şekil 3. Toplam puanların histogramı*

Analize dahil edilen tüm maddelerin p değerlerini bir arada grafikte görmek için *plot()* fonksiyonu kullanılır. *plot()* fonksiyonu içerisinde yer alan “ylim= c()” değeri y ekseninin maximum noktasını belirtir. "*main*” grafik başlığını, *“col”* çizgi rengini , “type” çizgi tipini , “*pch*” işaretçi şeklini, “*lwd*” çizgi genişliğini, “*xlab*” x ekseni etiketi, “*ylab*” y ekseni etiketini belirtmektedir. Aşağıda bazı grafik bilgilerine ait alternatif kodlar yer almaktadır:

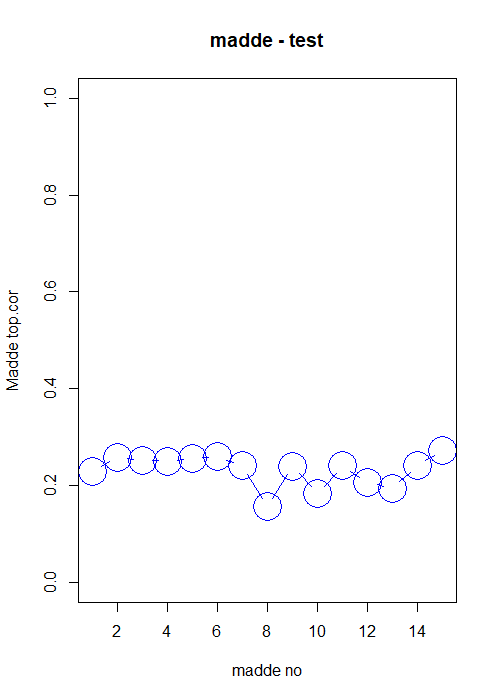
|  |
| --- |
| ##type alternatifleri:  "p"=nokta,  "b"=nokta ve cizgi,  "o"= nokta ve cizgi birarda  "l"= çizgi  # pch değeri 1-25 arasında farklı şekiller türetir. |

|  |
| --- |
| library(ggplot2)  library(dplyr)  #maddeler için p değeri grafiği  plot(DataOut [1, 1:15], ylim = c(0,1),  main= "Madde p value",  col="black",  type = "o",  pch=20,  lwd=5,  xlab="madde no”,  ylab="madde gucluk indeksi" ) |



*Şekil 4. Maddelerin güçlük düzeyleri*

|  |
| --- |
| # plot of item discrimination/difficulty  item.discrimination<-DataOut[6,]  item.difficulty<-DataOut[1,]  plot(item.discrimination,  type = "b",  pch = 1,  cex = 4,  col = "blue",  ylab = "Madde top.cor",  xlab = "madde no",  ylim = c(0, 1), xlim=c(1,15),  main = "madde - test") |

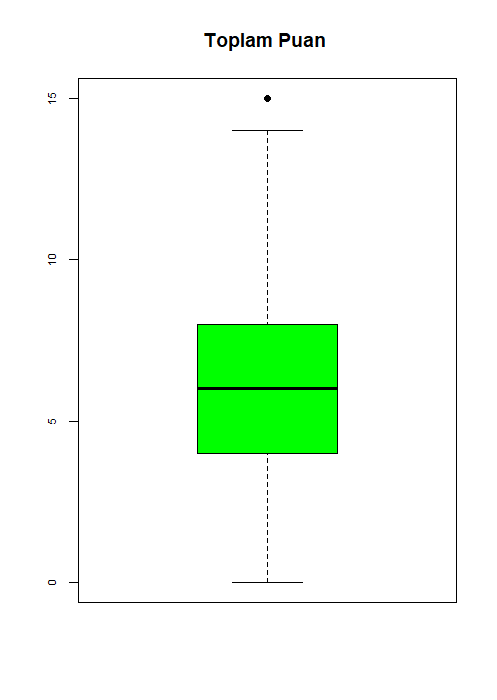


*Şekil 5. Madde test toplam korelasyonu*

Şekil 5’te yer alan madde – toplam test korelasyonu grafiği incelendiğinde maddelerin ayırt edicilik değerlerinin 0.20 düzeylerinde düşük seviyede olduğu yorumu yapılabilir. Ayrıca 8.maddenin en düşük korelasyona sahip olduğu görülmektedir. 0.20’den de küçük ayırt ediciliği olan 8. Maddenin testten çıkarılması gerekmektedir.

Madde toplam puanlarının boxplot ile gösterimi için gerekli komut ve çıktısı olan grafik aşağıda verilmiştir.

|  |
| --- |
| boxplot(Data[, 16],  main= "Toplam Puan ",  names = "Toplam Puan Degiskeni",  col= "green",  pch= 19,  cex.axis= 0.7) |

****

*Şekil 6. Toplam puanların boxplot grafiği*

Şekil 6’daki boxplotta toplam puanların 0 ile 15 puan aralığında değiştiği, 15 puana sahip tek bir birey olduğu, toplam puanların ortalamasının 6 puan civarında olduğu ve genel olarak normal dağılım gösterdiği söylenebilir.

Tüm maddelerin güçlük düzeylerinin grafiklerinin ayrı ayrı çizdirilmesi ve bu grafiklerin her birinin tek çıktı şeklinde görünmesi için *par(mfrow=c())* fonksiyonu kullanulabilir. 15 maddeye ait madde güçlük grafiklerini 5 satır ve 3 sütunda gösterebilmek için *par(mfrow=c(5,3))* komutu girilmiş ve ardından grafik çizdirme fonksiyonları bu komut ile tekrar çalıştırılmıştır. Çıktıya ait grafikler aşağıdaki şekilde sunulmuştur.

|  |
| --- |
| ok içeren bir resim  Açıklama otomatik olarak oluşturuldu |
| *Şekil 7. Tüm maddelere ait grafik çıktıları* |

## **Grafiklerin Kaydedilmesi**

R’da grafik kaydı dosya formatına göre uzantısı değişen fonksiyonlar ile kaydedilmektedir. Hangi dosya formatında kaydedilecekse dosya adının sonuna format bilgisi eklenir. Örneğin ; bmp(“*dosyaadı.bmp”),* jpeg(“*dosyaadı.jpg*”), pdf (“*dosyaadı.pdf*”), png (“*dosyaadı.png*”).

Grafik fonksiyonu öncesi istenilen dosya formatı kayıt fonksiyonu yazılır. Grafik kaydı *dev.off()* fonksiyonu kullanılarak bitirilir. Böylece grafikler ilgili dizinde belirtilen dosya formatı ile kaydedilir.

Bu çalışma kapsamında farklı formatlarda dosyalar oluşturularak çizdirilen grafikler *dev.off*() fonksiyonu ile kapatılmıştır. RStudio grafik alanında bu komutların çalışması ile bir grafik oluşturulmayacak , çizdirilen grafikler sadece çalışma dizininde oluşturulacak olan dosyaya yazdırılacaktır. Grafiğin dosyaya yazdırılmasında istenilen tür ve boyut değişikliği mümkündür.

|  |
| --- |
| ##grafiklerin kaydedilmesi  jpeg("Histogram.jpg")  plot( R1$pskkatg, R1$ismottoplam, pch = 19,col = "green")  dev.off()  ##png  ##2  #grafiği a4 boyutunda kaydetmek için  pdf("pskljk iklm kategori.pdf", paper="a4")  plot( R1$pskkatg, R1$ismottoplam, pch = 19,col = "green")  dev.off()  #grafiklerin dönglerle kaydedilmesi  #her bir dosyanın ayrı bir pdf dosyası olarak kaydedilmesi  #paste() fonksiyonu  for(i in 1:15) {  pdf(paste("ToplamPuanHistogramı", i, ".pdf", sep=""), paper= "a4")  hist(Data$X[Data$X == i], xlim = c(-4,4),  main=paste("X",i),  xlab="ToplamPuan")  dev.off()  }  #grafiklerin tek bir dosyaya kaydedilmesi  pdf("ToplamPuanlaraGöreHistogramlar.pdf", paper="a4")  for(i in 0:15){  hist(Data$X[Data$X==i], xlim=c(-4,4),  main=paste("ToplamPuan",i),  xlab="Toplam puan")  }  dev.off() |

Tablo 1. Testin KTK’ya göre madde indeksleri.

|  |  |  |
| --- | --- | --- |
| Madde | Madde Güçlüğü | Ayırt Edicilik İndeksi |
| 1 | 0,571 | 0,230 |
| 2 | 0,479 | 0,259 |
| 3 | 0,447 | 0,252 |
| 4 | 0,435 | 0,251 |
| 5 | 0,420 | 0,256 |
| 6 | 0,389 | 0,260 |
| 7 | 0,365 | 0,241 |
| 8 | 0,347 | 0,157 |
| 9 | 0,295 | 0,240 |
| 10 | 0,235 | 0,184 |
| 11 | 0,232 | 0,241 |
| 12 | 0,611 | 0,206 |
| 13 | 0,617 | 0,195 |
| 14 | 0,421 | 0,241 |
| 15 | 0,433 | 0,272 |
| Ortalama | 6,297 |  |
| Standart sapma | 2,811 |  |
| Güvenirlik | 0,603 |  |

1000 kişinin yanıtladığı 21 sorudan rastgele seçilen ve kodlanan 15 maddeden oluşan testin madde güçlüğü 0.23 ile 0.617 arasında değişmektedir. Testin ortalama güçlüğü 0.419, standart sapması 2,811 olarak hesaplanmıştır. Testteki maddelerin güvenirliği 0.22 ile 0.29 arasında değişmektedir. Testin güvenirliği 0.603 olarak hesaplanmıştır.

# KAYNAKÇA

Atar, B. ,Atalay, Kabasakal , K. vd. (2021). *R ile veri analizi ve psikometri uygulamaları* (4.Baskı).Pegem.

Büyüköztürk, Ş., Çokluk, Ö. & Köklü, N. (2016). *Sosyal bilimler için istatistik.* Pegem Akademi.

<https://www.geeksforgeeks.org/data-visualization-with-r-and-ggplot2/?ref=lbp>

<https://www.learnr4free.com/tr/index.html>

<https://r-graph-gallery.com/48-grouped-barplot-with-ggplot2>

<https://rpubs.com/hmztnc/461183>

https://r4ds.had.co.nz/

Şirin,S.(2018). *R ile Uygulamalı Analiz yöntemleri. “*<https://www.researchgate.net/profile/SelahattinSirin/publication/325930147_R_ile_Uygulamali_Analiz_Yontemleri_-_I/links/5f3252a8a6fdcccc43bf0b53/R-ile-Uygulamali-Analiz-Yoentemleri-I.pdf>” adresinden erişilmiştir.