

whyR Eğitim Bilimleri Geliştirilen R Paketleri

Daha ileriye... En iyiye...

Dr. Kübra ATALAY KABASAKAL



**HACETTEPE
ÜNİVERSİTESİ**

www.hacettepe.edu.tr



Eğitim Bilimleri Geliştirilen R Paketleri

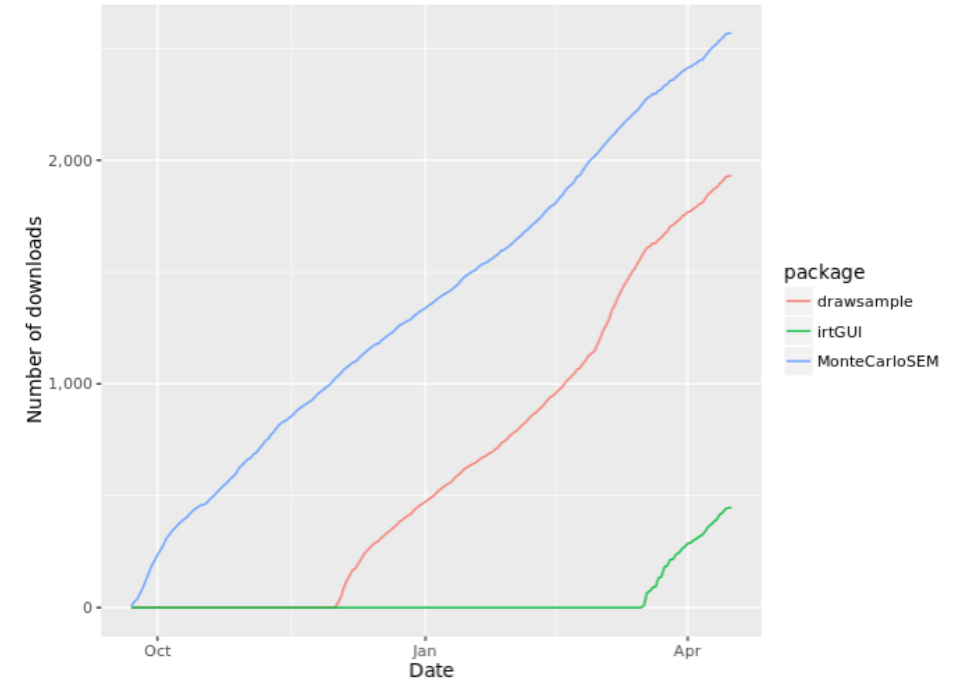
Kalsik Test Kuramı	Madde Tepki Kuramı	DIF-Eşitleme-CAT	Faktör Analizi-Yapısal Eşitlik Modelleri
CTT	eRm	kequate	lavaan
CTTShiny	ltm	equateIRT	SEM
psych	mirt	difR	psych
subscore	TAM	catR	semPlot
		mirtCAT	MplusAutomation
		xxIRT	lisreltoR

Türkler tarafından geliştirilen psikometri alanında kullanılan R paketleri

```
# install.packages("itemanalysis", "profileR", "eirm", "hemp",  
                  "MonteCarloSEM", "mlmmm", "irtGUI", "drawsample")
```

```
library(itemanalysis)  
library(profileR)  
library(eirm)  
library(hemp)  
  
library(mlmmm)
```

```
library(MonteCarloSEM)  
library(irtGUI)  
library(drawsample)
```



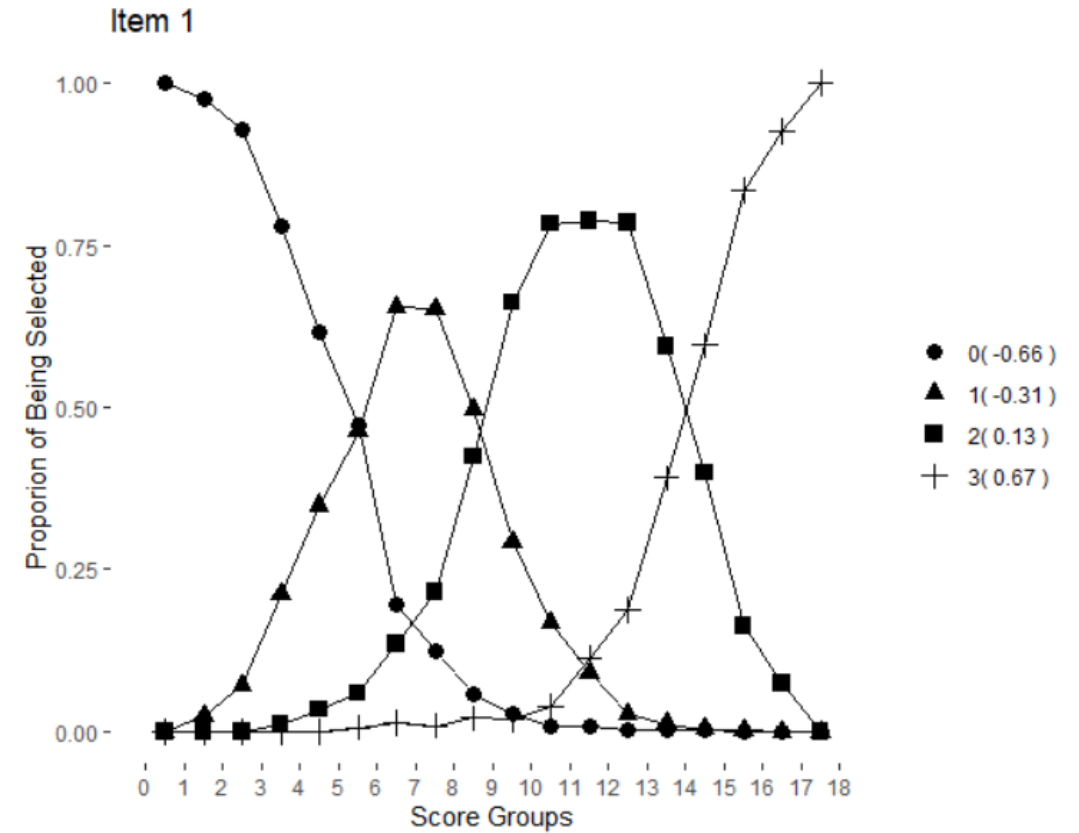
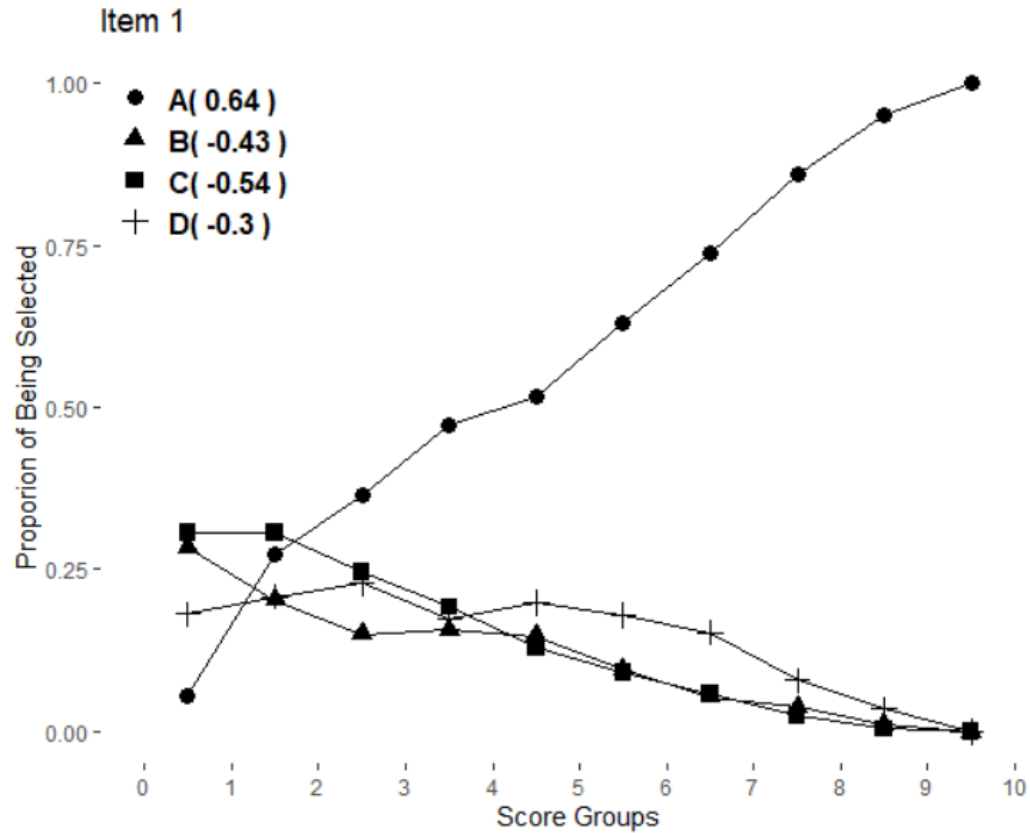
itemanalysis

Paket Adı	Tanıtımı	Geliştirici/leri - Tarih- İletişim adresleri
itemanalysis	Klasik Test Kuramına göre madde analizi	Cengiz Zopluoglu - 2018 - (c.zopluoglu@miami.edu2)

- Klasik Test Kuramına göre çoktan seçmeli maddeler ve çok kategorili puanlanan maddeler için madde istatistikleri hesaplayan bu pakette iki temel fonksiyon (`itemanalysis1`, `itemanalysis2`) ve bu fonksiyonların çalışmasında kullanılabilecek iki veri seti bulunmaktadır.
- Paketi CTT paketinden farklı kılan çok kategorili maddeler için madde istatistikleri hesaplayabilmesi, çeldiriciler için cevaplanma oranı, çeldiriciler için nokta çift-serili ve çift-serili korelasyon hesabı yapmasıdır. Ayrıca her bir celdirinin kategorileri ve toplam puana ilişkin item trace line grafikleri sunmaktadır.



itemanalysis





profileR

Paket Adı	Tanıtımı	Geliştirici/leri - Tarih- İletişim adresleri
profileR	Çok Değişkenli Yöntemler için Profil Analizi	Okan Bulut & Christopher D. Desjardins - 2018 - (okanbulut84@gmail.com, cddesjardins@gmail.com)

Bulut, O., & Desjardins, C. D. (2020). *Profile Analysis of Multivariate Data: A Brief Introduction to the profileR Package*. VV(ii). <https://doi.org/10.31234/osf.io/sgy8m>

profileR

- Eğitimsel veya psikolojik bir ölçümden elde edilen test sonuçları ile ilgili olan profil analizi, MANOVA'nın tekrarlı ölçümlerde bir eşdeğeri olarak görülebilir.
- Eğitimsel testlerde puan profilleri, öğrencilerin güçlü ve zayıf yanlarını belirleyip, onların yönlendirilmesinden yardımcı olabilir. Psikolojik ölçümlerde ise, hangi özelliklerin veya faktörlerinin bir kişinin gizil bir yapı üzerindeki klinik profilinin modelinde bir dereceye kadar kendini gösterdiğini belirlemek için bir test puan profili kullanılabilir.
- ProfileR paketi hem bir bireyin alt testlerdeki genel performansı hem de alt testlerdeki puanlar arasındaki değişimi ortaya çıkaracak analizlere imkan sağlamaktadır.
- Pakette kriterle ilişkili profil analizi *cpa()*, çok boyutlu ölçekleme ile profil analizi *pams()*, gruplara göre profil analizi *pbg()*, birey içi faktör modeline *wprifm()* fonksiyonları, profil güvenilirliği belirlemek amacıyla *pr()* fonksiyonu ve örnek veri setleri yer almaktadır. Bunlar yanında çeşitli görselleştirme olanakları da sunmaktadır.

eirm

Paket Adı	Tanıtımı	Geliştirici/leri - Tarih- İletişim adresleri
eirm	İki ve Çok kategorili Maddeler İçin Açıklayıcı Madde Tepki Kuramı	Okan Bulut - 2021- (bulut@ualberta.ca)

Paket, açıklayıcı madde tepki modellemesini (Boeck ve Wilson, 2004) kullanarak iki ve çok kategorili maddeleri analiz etmek ve görselleştirmek için farklı fonksiyonlar sağlamaktadır.

eirm

- Paketin temel fonksiyonu *eirm()*, MTK modellerine bireylerin ve maddelerin özelliklerini eklemeye ve bunların etkilerini tek adımda incelemeye imkân sağlamaktadır.
- Paketin temel fonksiyonunda “response ~ -1 + predictor1 + predictor2” şeklinde regresyon modeline benzer bir formül ile sağlanır.
- Pakette çok kategorili madde cevaplarını uzun formata getirebilmek için *polyreformat()* fonksiyonu bulunmaktadır. Paket içerisinde çıktıları elde etmek amacıyla *print.eirm()* fonksiyonu, birey-madde haritalarını elde etmek için ise *plot.eirm()* fonksiyonu yer almaktadır.
- Ayrıca *eirmShiny()* isimli fonksiyon ile paket fonksiyonlarına ara yüz sağlamaktadır.



hemp

Paket Adı	Tanıtımı	Geliştirici/leri - Tarih- İletişim adresleri
hemp	Handbook of Educational Measurement and Psychometrics Using R	Christopher D. Desjardins - Okan Bulut - 2018

Desjardins & Bulut (2018) kitaplarının eşlikçisi olarak hazırlanmış olan paket, kitabın bölümlerinde kullanılan veri setleri ve fonksiyonları içermektedir.

MonteCarloSEM

Paket Adı	Tanıtımı	Geliştirici/leri - Tarih- İletişim adresleri
MonteCarloSEM	Monte Carlo Data Simulasyon	Fatih Orcan - 2020 - (fatihorcan84@gmail.com)

Paket farklı koşullar için yapısal eşitlik modellerine uygun simülatif veriler üretip analiz edebilmektedir.

Pakette yer alan iki temel fonksiyonundan biri olan *sim.normal()* fonksiyonu ile oluşturulması istenen veri seti sayısı, örneklem büyüklüğü, faktör yükleri ve faktörler arası korelasyon matrisi girdileri ile istenilen dizini üretilmiş veriler aktarılabilir.

MonteCarloSEM

- Fonskiyon çalıştırıldığında dizine üretilen verilerin yanında faktor korelasyon ve faktor yüklerini bilgileri “Model_Info.dat” dosyasına, üretilen dosya adları ise “Data_List.dat” adlı dosyaya akatrılır.
- *simnormal()* fonksiyonun girdileri olan faktörler arası korelasyon matrisi *fcors.value()* fonksiyonu ile, faktor yükleri matrisi ise *loading.value()* fonksiyonu ile oluşturulabilir.
- *sim.skewed()* fonksiyonu ise Fleishman’ın güç yöntemini kullanarak istenilen madde dağılımların çarpık oluşturulmasını sağlanabilir. Bu amaçlar fonksiyon içinde nonnormal ve Fleishman argümanlarının kullanılmasını gerektirmektedir.
- *fit.simulation()* fonksiyonu ile ise oluşturulan model ve oluşturuması istenen modelle uyumu arka planda lavaan paketi çalıştırılarak uyum indeksleri ile değerlendirebilmektedir.



irtGUI

Paket Adı	Tanıtımı	Geliştirici/leri - Tarih- İletişim adresleri
irtGUI	Madde Tepki Kuramı Analizleri Uygulaması	Huseyin Yildiz - 2021 - (huseyinyildiz35@gmail.com)

MTK'ya ait model veri uyumu sınaması, madde ve yetenek parametreleri kestirimleri, boyutluluk varsayımının sınanması, madde karakteristik, madde bilgi ve kategori karakteristik eğrileri elde edilebilir. Bu işlevler Tek Boyutlu MTK modellerinden iki kategorili maddeler için Rasch, 2 PLM, 3 PLM, çok kategorili maddelerin için ise Aşamalı Tepki Modeli ve Genelleştirilmiş Kısmi Puan Modelleri altında kullanılabilmektedir.

drawsample

Paket Adı	Tanıtımı	Geliştirici/leri - Tarih- İletişim adresleri
drawsample	Orneklem Secme	Kubra Atalay Kabsakal - 2021- (katalay@hacettepe.edu.tr)

“drawsample” paketi gerçek veri setlerinden istenilen özelliklerde veri setleri çekebilmek amacıyla geliştirilmiştir. Paket ile seçilecek örneklemin sahip olması beklenen özelliklerden normallikten sapma ölçüleri (çarpıklık ve basıklık) ve örneklem büyüklüğü gibi koşulların bir veya birden fazlasının belirlenmesi ile kasıtlı örnekleme yapılabilmektedir.

drawsample

- Simülasyon istenilen dağılım özelliklerinde veriler elde edilmesinde güçlü bir teknik olmasına rağmen, gerçek verilerde ulaşılamayan mükemmel bir uyum sağlaması simülasyon verilerinin kullanılmasında bazı sınırlılıklar oluşturmaktadır (Wicklin, 2013).
- Halgren (2013) gerçek veri setlerinin, idealist koşullar altında üretilen "temiz" simülasyon veri setlerinden daha "kirli" olacağını,
- Sireci (1991) ise simülatif veriler kullanıldığında, üretilen verilerin pratikte karşılaşılan ilgili durumun özelliklerini doğru olarak yansıtip yansıtmadığı bilinemeyeceğini ve geçerliğinin test edilemeyeceğini ifade etmiştir.
- Ayrıca, Educational Measurement: Issues and Practice (EM: IP) ve Journal of Educational and Behavioral Statistics (JEBS) gibi bazı prestijli dergiler, uzun bir süre zarfında simülasyon çalışmalarını kabul etseler de günümüzde simülasyon temelli çalışmaların "uygun olmayan makale konu örneklerinden" veya "düşük önceliğe sahip" olarak kabul edildiğini belirtmiştir (John Wiley & Sons Inc., 2019; American Educational Research Association, 2020).

drawsample

- Gerçek uygulamalarda veri toplama süreci zorluklarla doludur.
- Elde edilen örneklem evren dağılımını temsil etmiyor, normal dağılmıyor veya istenen dağılıma uygun olmayan bir halde olabilir.
- Araştırma problemlerine dayalı olarak normalliği sağlamada, normal dağılımdan uzaklaştıran verilerin çıkarıldığı ya da ayrıştırıldığı bazı araştırmalara rastlanmıştır. Literatürde birçok araştırmacı, yaptıkları çalışmaların amacına uygun olacak şekilde gerçek veri setinden (evren) örneklem almayı seçmiştir (Courville, 2004; Çelikten ve Çakan, 2019, Doğan ve Kılıç, 2018; Fan, 1998; Nartgün, 2002; Reyhanlıoğlu Keçeoğlu, 2018).
- Araştırmalar incelendiğinde, araştırmacıların geniş bir veri setinden istenen özelliklere sahip örneklem seçimini sağlayacak bir araca ihtiyaç olduğu düşünülerek bu paket geliştirilmiştir.

drawsample

- *draw_sample()*

Description

A function to sample data with desired properties.

Usage

```
draw_sample(  
  dist,  
  n,  
  skew,  
  kurts,  
  replacement = FALSE,  
  output_name = c("sample", "default")  
)
```

- *draw_sample_n()*

Description

A function to sample data with desired properties.

Usage

```
draw_sample_n(  
  dist,  
  n,  
  skew,  
  kurts,  
  location = 0,  
  delta_var = 0,  
  output_name = c("sample", "default")  
)
```



<i>Skewness</i>	<i>Kurtosis (min)</i>	<i>Kurtosis (max)</i>	<i>Skewness</i>	<i>Kurtosis (min)</i>	<i>Kurtosis (max)</i>
0	-1.2	20	1.9	4.4	20
0.1	-1.2	20	2	5	20
0.2	-1.1	20	2.1	5.6	20
0.3	-1.1	20	2.2	6.3	20
0.4	-0.9	20	2.3	7.1	20
0.5	-0.8	20	2.4	7.8	20
0.6	-0.6	20	2.5	8.6	20
0.7	-0.4	20	2.6	9.5	20
0.8	-0.2	20	2.7	10.4	20
0.9	0.1	20	2.8	11.4	20
1	0.4	20	2.9	12.4	20
1.1	0.7	20	3	13.4	20
1.2	1	20	3.1	14.4	20
1.3	1.4	20	3.2	15.5	20
1.4	1.8	20	3.3	16.5	20
1.5	2.3	20	3.4	17.6	20
1.6	2.7	20	3.5	18.8	20
1.7	3.2	20	3.6	19.9	20
1.8	3.8	20			

drawsample

```
> constants_table
# A tibble: 5,292 x 5
  Skew Kurtosis      b      c      d
  <dbl>   <dbl> <dbl>   <dbl>   <dbl>
1      0    -1.2  1.36  1.60e-14 -0.136
2      0    -1.1  1.28  7.82e-12 -0.103
3      0     -1  1.22  1.73e-10 -0.0802
4      0    -0.9  1.18  2.93e-10 -0.0643
5      0    -0.8  1.15 -1.30e-10 -0.0520
6      0    -0.7  1.12  2.39e-12 -0.0420
7      0    -0.6  1.10  1.23e- 8 -0.0335
8      0    -0.5  1.08  1.04e- 9 -0.0263
9      0    -0.4  1.06  8.65e-11 -0.0199
10     0    -0.3  1.48  2.99e-14 -0.206
# ... with 5,282 more rows
```

Fleishman'ın (1978) güç yöntemi, istenen dağılım özelliklerine sahip bir örnek elde etmek için kullanılır.

Fleishman (1978), önceden belirlenmiş momentlerle bir dağılım oluşturmak için standart bir normal değişkenin kübik dönüşümünü kullandı.

drawsample

```
if (skew < 0) {  
  constants_table$c <- -1 * constants_table$c  
  constants_table$Skew <- -1 * constants_table$Skew  
}  
if (skew %in% constants_table$Skew == FALSE) {  
  stop("No valid power method constants could be found for\    the specified values. Change the values")  
}  
else if (skew %in% constants_table$Skew == TRUE & kurts %in%  
  constants_table[constants_table$Skew == skew, ]$Kurtosis ==  
  FALSE) {  
  stop("No valid power method constants could be found for\    the specified values. Change the values")  
}
```

Bu kısımda kişinin girdiği çarpıklık ve basıklık değerleri control ediliyor

drawsample

```
# conduct Fleishman's power method for the specified
# skewness and standardized kurtosis
repeat {
  for (i in 1:dim(constants_table)[1]) {
    reference <- stats::rnorm(n, 0, 1)
    constants <- constants_table[i, 3:5]
    b <- constants$b
    c <- constants$c
    d <- constants$d
    reference_v2 <- -c + b * reference + c * (reference^2) +
      d * (reference^3)
    skew_value <- round(psych::describe(reference_v2)$skew,
      1)
    kurt_value <- round(psych::describe(reference_v2)$kurtosis,
      1)
    if (skew_value == skew & kurt_value == kurts) {
      reference_v3 <- reference_v2
      break
    }
  }
  if (is.null(reference_v3) == FALSE)
    break
}
```

Bu kısımda ise Fleishman'ın güç yöntemi uygulanarak oluşan yeni dağılımın çarpıklık ve basıklık dağılımı Kullanıcının girdiği ile karşılaştırılır.

drawsample

```
reference_v4 <- scale_ref(reference_v3, from = min(x), to = max(x))
x_counts <- graphics::hist(reference_v4)$counts
n_break <- length(graphics::hist(reference_v4)$breaks) -
  1
x_break <- graphics::hist(reference_v4)$breaks
x_v1 <- as.numeric(cut(x, x_break, include.lowest = TRUE))
dist2 <- data.frame(dist, x_v1)
x_n <- unname(table(x_v1))
control <- sum(x_n >= x_counts)
if (control != length(x_counts)) {
  if (replacement == FALSE) {
    stop("Cannot take a sample form that data without replacement.\n
      Please change replacement=TRUE")
  }
}
```

Bu bölümde hem referans dağılımı hem de asıl dağılım kesikli hale getirilir.

Oluşan kategorilerde eğer herhangi bir kategoride frekans sayısı, asıl dağılımının ilgili kategorisinin frekansı, referans dağılımının ilgili kategorisinin frekansı geçmiyorsa yeniden örnekleme yapılması gerektiği ile ilgili uyarı ile fonksiyon durdurulur

drawsample

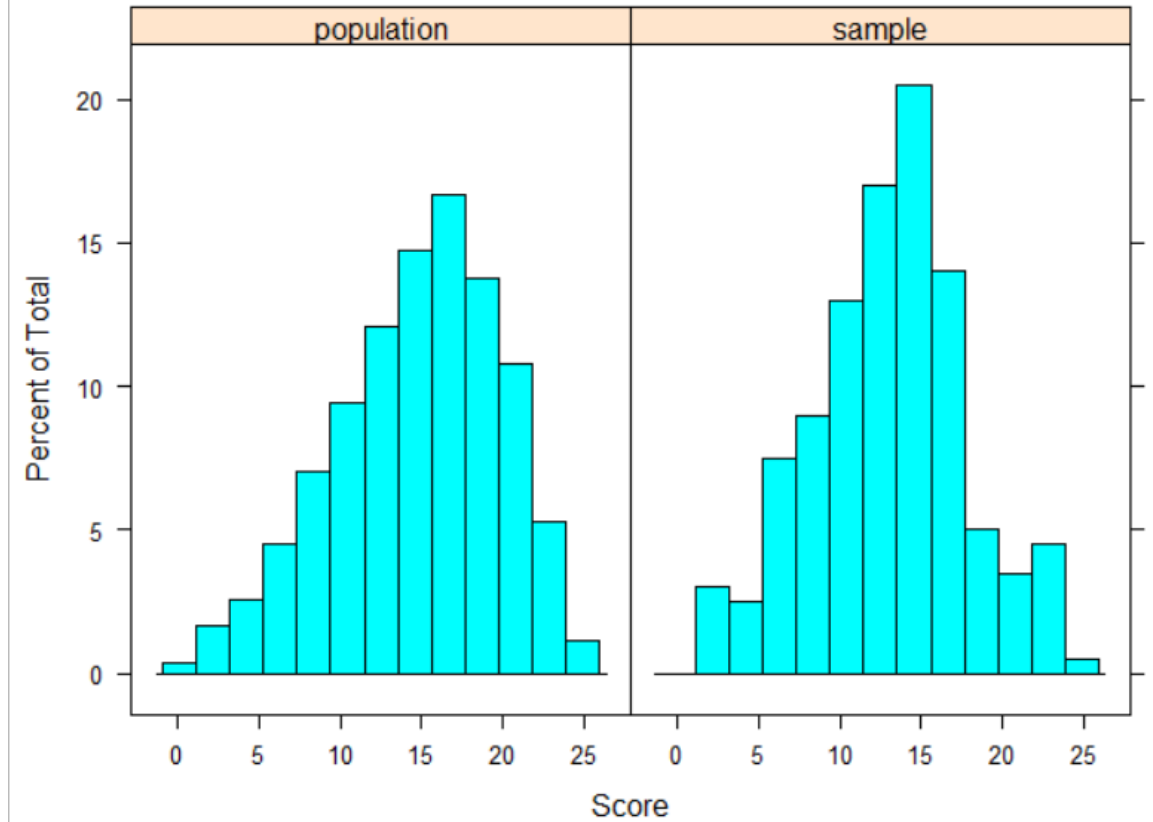
```
new_sample <- matrix(NA, nrow = n_break, ncol = max(x_counts))
ID_list <- matrix(NA, nrow = n_break, ncol = max(x_counts))
new_sample_2 <- list()
ID_list_2 <- list()
for (i in 1:n_break) {
  new_count <- 0
  j <- 0
  while (new_count < x_counts[i]) {
    j <- j + 1
    IDx <- dplyr::filter(dist2, x_v1 == i)
    IDx <- dplyr::sample_n(IDx, 1)
    if (replacement == FALSE) {
      dist2 <- dplyr::filter(dist2, id != IDx$id)
    }
    else {
      dist2 <- dist2
    }
    new_count <- new_count + 1
    new_sample[i, j] <- IDx$x
    ID_list[i, j] <- IDx$id
  }
}
```

Bu kısımda ise yeni veri seti oluşturuluyor.



drawsample

```
> ornek1 <- draw_sample(dist=example_data[,c(1,2)],n=200,skew = 0,kurts = 0,  
+ output_name = c("sample", "1"))  
> ornek1  
$desc  
      n mean  sd min max skew kurtosis  
population 5000 14.61 4.90 0 25 -0.40 -0.35  
reference 200 12.35 4.47 0 25 0.00 0.02  
sample 200 12.93 4.57 2 25 -0.03 -0.10  
  
$sample  
# A tibble: 200 x 2  
   id      x  
  <dbl> <dbl>  
1  1165     2  
2  2187     2  
3   605     4  
4  4752     3  
5  2871     3  
6  1486     3  
7  4212     3  
8  2457     5  
9  4119     6  
10  886      6  
# ... with 190 more rows
```



drawsample

- Fonksiyon kullanımında, gerçek veri setinden istenen özelliklere mükemmel bir şekilde uyan örneklem seçmenin, özellikle yeniden örnekleme yapılmadığında, kolay olmadığı unutulmamalıdır.
- Fonksiyonun tutarlılığının değerlendirilmesi amacıyla üretilen simülasyon verisinden tekrarlı çekilen örneklemelere dayalı olarak *draw_sample()* fonksiyonunun, daha az rastlanan büyük çarpıklık değerlerinde daha tutarsız sonuçlar verdiği görülmüştür. Ayrıca bu tutarsızlık, örneklemin çekildiği evrenin özellikleriyle de doğrudan ilişkilidir.
- Evren ile çekilecek örneklemin büyüklükleri ve istenen özelliklere sahip örneklemin evren ile dağılım türlerinin benzerliği tutarsızlık miktarı ile doğrudan ilişkilidir.
- Rastgele atamanın doğası gereği, fonksiyondaki argümanların aynı değerleri için bile her seferinde farklı örneklem çekebilir. Kullanıcılar istenilen özelliklere sahip örnekleme ilk seferde elde edemezler ise kullanıcılara fonksiyonu birkaç kez çalıştırmaları önerilir.



drawsample

- Atalay Kabasakal, K. & Gündüz, T. (2020). Drawing a sample with desired properties from population in R package “drawsample”. *Journal of Measurement and Evaluation in Education and Psychology*, 11(4), 405-429. doi: 10.21031/epod.790449



**HACETTEPE
ÜNİVERSİTESİ**

www.hacettepe.edu.tr

Teşekkürler