Parametrik Olmayan İstatistiksel Yöntemler Ödevi

Ramazan Erduran, İlkay Şafak Baytar

04 Aralık 2021

### (1) Bağımsız veri setine ilişkin inceleme:

Verilerin alındığı kaynak olarak [Datacamp](datacamp.com) seçildi. İlk olarak veriler gapminder\_tsv beklenen ömür değişkeni ile filtrelendi gapminder\_lifeExp.

Daha sonra bu filtreler üç ülkeye **Europe**, **Americas**, **Asia** bölündü ki daha rahat ve temiz analiz yapabilelim.

gapminder\_lifeExp <- gapminder\_tsv %>%  
 filter(continent %in% c("Asia", "Europe", "Americas")) %>%  
 summarize(continent=continent, lifeExp=lifeExp) %>%  
 arrange(desc(continent))  
  
gp\_eu <- gapminder\_lifeExp %>%  
 filter(continent=="Europe")  
  
gp\_usa <- gapminder\_lifeExp %>%  
 filter(continent=="Americas")  
  
gp\_asia <- gapminder\_lifeExp %>%  
 filter(continent=="Asia")  
#Kitle:  
head(gapminder\_lifeExp)

## continent lifeExp  
## 1 Europe 55.23  
## 2 Europe 59.28  
## 3 Europe 64.82  
## 4 Europe 66.22  
## 5 Europe 67.69  
## 6 Europe 68.93

#İncelenecekler:  
head(gp\_eu)

## continent lifeExp  
## 1 Europe 55.23  
## 2 Europe 59.28  
## 3 Europe 64.82  
## 4 Europe 66.22  
## 5 Europe 67.69  
## 6 Europe 68.93

head(gp\_asia)

## continent lifeExp  
## 1 Asia 28.801  
## 2 Asia 30.332  
## 3 Asia 31.997  
## 4 Asia 34.02  
## 5 Asia 36.088  
## 6 Asia 38.438

head(gp\_usa)

## continent lifeExp  
## 1 Americas 62.485  
## 2 Americas 64.399  
## 3 Americas 65.142  
## 4 Americas 65.634  
## 5 Americas 67.065  
## 6 Americas 68.481

#Örneklem çekme:

Analize başlamak üzere n=30 için 3 değişkenden de veriler rastgele çekildi ve bu örneklemler: gp\_eu\_samp, gp\_asia\_samp, gp\_usa\_samp olarak adlandırıldı.

as.numeric kodunu kullandık ki çektiğimiz örneklemler “character” formatında kalmasın ve analizini yapabilelim.

#Tanımlayıcı istatistikler ve normallik varsayımı testi:

Özetleyici istatistikler kapsamında incelenecek her bir grup için stat.desc kodu kullanıldı.

Normallik varsayımında ise: H0: Verilerin dağılımı ile normal dağılım arasında fark yoktur. H1: Verilerin dağılımı ile normal dağılım arasında fark vardır. şeklinde iki hipotez kurulup %95 güven düzeyinde (alpha=0.05) test edildi.

Test süresince (n=30) < 50 için **Shapiro Wilk** testi kullanıldı. Test sonucunda ise her bir grup için p-value < 0.05 olduğundan mütevellit; “%95 güven düzeyiyle test edilecek grupların dağılımı ile normal dağılım arasında fark vardır” yorumu yapılabilir.

# Europe'a ait özetleyici istatistikler:  
stat.desc(gp\_eu\_x$lifeExp)

## nbr.val nbr.null nbr.na min max range   
## 3.000000e+01 0.000000e+00 0.000000e+00 4.807900e+01 7.853000e+01 3.045100e+01   
## sum median mean SE.mean CI.mean.0.95 var   
## 2.144654e+03 7.224500e+01 7.148847e+01 1.111415e+00 2.273098e+00 3.705727e+01   
## std.dev coef.var   
## 6.087469e+00 8.515316e-02

#Asia'ya ait özetleyici istatistikler:  
stat.desc(gp\_asia\_x$lifeExp)

## nbr.val nbr.null nbr.na min max range   
## 30.0000000 0.0000000 0.0000000 31.9970000 78.7700000 46.7730000   
## sum median mean SE.mean CI.mean.0.95 var   
## 1891.7630000 64.2300000 63.0587667 1.9170334 3.9207736 110.2505171   
## std.dev coef.var   
## 10.5000246 0.1665117

#Americas'a ait özetleyici istatistikler:  
stat.desc(gp\_usa\_x$lifeExp)

## nbr.val nbr.null nbr.na min max range   
## 30.0000000 0.0000000 0.0000000 37.5790000 77.8600000 40.2810000   
## sum median mean SE.mean CI.mean.0.95 var   
## 2014.9420000 68.5745000 67.1647333 1.4993550 3.0665253 67.4419637   
## std.dev coef.var   
## 8.2123056 0.1222711

#Normallik testi:  
shapiro.test(gp\_eu\_x$lifeExp)

##   
## Shapiro-Wilk normality test  
##   
## data: gp\_eu\_x$lifeExp  
## W = 0.81509, p-value = 0.000125

shapiro.test(gp\_asia\_x$lifeExp)

##   
## Shapiro-Wilk normality test  
##   
## data: gp\_asia\_x$lifeExp  
## W = 0.91877, p-value = 0.02492

shapiro.test(gp\_usa\_x$lifeExp)

##   
## Shapiro-Wilk normality test  
##   
## data: gp\_usa\_x$lifeExp  
## W = 0.83329, p-value = 0.0002825

# p-value < alpha=0.05 için %95 güven düzeyiyle söylenebilir ki veriler ile normal dağılım arasında fark vardır.