

Laboratorium 3

1. Test jednorodności wariancji

Oprogramować jednostajnie najmocniejszy test wśród testów nieobciążonych dla problemu testowania równości wariancji w modelu Gaussowskim. Przetestować jednorodność wariancji na danych z Baza 2.2.

```
TV<-function(X,Y)

{

  nx<-length(X)

  ny<-length(Y)

  T<- sum((Y-mean(Y))^2)/ (sum((Y-mean(Y))^2)+ sum((X-mean(X))^2))

  F<-function(x,t)

  {

    pbeta(x,(ny-1)/2,(nx-1)/2)- pbeta(x,(ny+1)/2,(nx-1)/2)- pbeta(t,(ny-1)/2,(nx-1)/2)+ pbeta(t,(ny+1)/2,(nx-1)/2)

  }

  if(T<((ny-1)/(nx+ny-2)))

  {

    a<-c()

    b<-c()

    a[1]= ((ny-1)/(nx+ny-2))

    b[1]=1

    for (i in 1:40)

    {

      d<-(a[i]+b[i])/2

      if(F(d,T)<0)

      {

        a[i+1]=a[i]

        b[i+1]=d

      }

      else

      {

        a[i+1]=d

        b[i+1]=b[i]

      }

    }

    d=(a[40]+b[40])/2

    p<-pbeta(T,(ny-1)/2,(nx-1)/2)+1-pbeta(d,(ny-1)/2,(nx-1)/2)
```

```

    }

    if(T>=((ny-1)/(nx+ny-2)))

    {

      a<-c()

      b<-c()

      a[1]=0

      b[1]==((ny-1)/(nx+ny-2))

      for (i in 1:40)

      {

        d<-(a[i]+b[i])/2

        if(F(d,T)<0)

        {

          a[i+1]=d

          b[i+1]=b[i]

        }

        else

        {

          a[i+1]=a[i]

          b[i+1]=d

        }

      }

      d=(a[40]+b[40])/2

      p<-pbeta(d,(ny-1)/2,(nx-1)/2)+1-pbeta(T,(ny-1)/2,(nx-1)/2)

    }

    return(p)

  }

```

```

library(readxl)

Baza<- read_excel("Baza 2.2.xlsx")

X<-na.exclude(Baza$"Zdrowi")

Y<-na.exclude(Baza$"Chorzy")

TV(X,Y)

```

Narysować wykres empirycznej mocy powyższego testu dla 20 obserwacji X z rozkładu $N(2,10)$ i 20 obserwacji Y z rozkładu $N(5,s)$ gdzie $s \in [1,25]$.

```
moc<-c()

m<-c()

mc<-1000

s<-seq(1,25,by=1)

for (k in 1:25)
{
  for (j in 1:mc)
  {
    X<-rnorm(20,2,10)
    Y<-rnorm(20,5,s[k])
    p<-TV(X,Y)
    m[j]=0
    if(p<0.05){m[j]=1}
  }
  moc[k]<-mean(m)
}

plot(moc)
```

Gotowy test jednorodności wariancji (F-test) dla wektorów X i Y uzyskujemy formułą:

```
var.test(X,Y)
```

Nie jest to jednak test jednostajnie najmocniejszy wśród nieobciążonych. Oparty jest on na statystyce testowej Z opisanej na wykładzie jednak do obliczenia p-wartości nie używa wartości spełniających założenia twierdzenia o teście jednostajnie najmocniejszym wśród nieobciążonych (po prostu podwaja pole na lewo lub na prawo od statystyki testowej w zależności od jej wartości pod gęstością odpowiedniego rozkładu Snedecora-Fishera czyli odcina ogony o tym samym prawdopodobieństwie).

Zadanie do samodzielnego wykonania: Sporządzić na jednym rysunku wykres empirycznej mocy dla F-testu i testu jednostajnie najmocniejszego wśród nieobciążonych dla 30 obserwacji X z rozkładu $N(2,10)$ i 70 obserwacji Y z rozkładu $N(5,s)$ gdzie $s \in [1,25]$. Czy wykresy różnią się znacząco od siebie?

```
moc1<-c()

m1<-c()

moc2<-c()

m2<-c()

mc<-1000

s<-seq(1,25,by=1)
```

```

for (k in 1:25)

{

  for (j in 1:mc)

  {

    X<-rmnorm(30,2,10)

    Y<-rmnorm(70,5,s[k])

    p1<-TV(X,Y)

    p2<-var.test(X,Y)$p.value

    m1[j]=0

    m2[j]=0

    if(p1<0.05){m1[j]=1}

    if(p2<0.05){m2[j]=1}


  }

  moc1[k]<-mean(m1)

  moc2[k]<-mean(m2)

}

plot(moc1,lwd=2)

lines(moc2, col="blue", lwd=2)

```

Zadanie do samodzielnego wykonania: pokazać obciążenie testu F (patrz **Laboratorium 3 rys**)

```

moc2<-c()

m2<-c()

mc<-1000000

r<-seq(0,1,by=0.1)

for (k in 1:11)

{

  for (j in 1:mc)

  {

    X<-rnorm(10,2,10)

    Y<-rnorm(50,5,10+(k-1)/10)

    p2<-var.test(X,Y)$p.value

    m2[j]=0

    if(p2<0.05){m2[j]=1}

  }

  moc2[k]<-mean(m2)

```

```
}
```

```
plot(moc2,r)
```