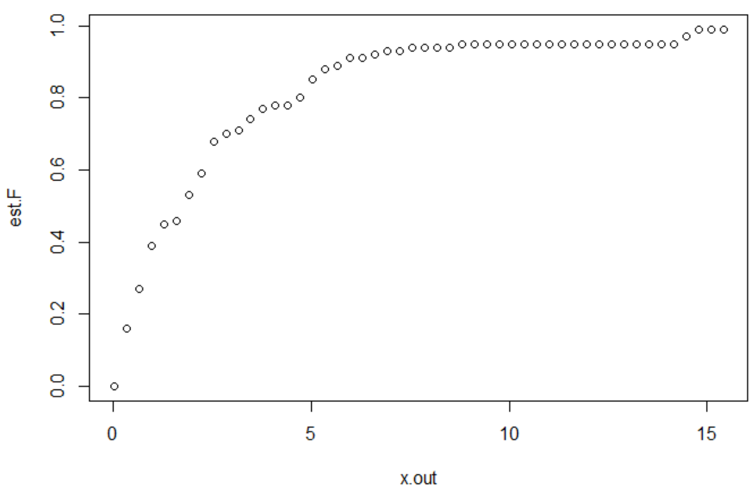
1. A) Παρουσιάστε διάγραμμα της εκτιμώμενης αθροιστικής συνάρτησης κατανομής F

   Η εμπειρική ασκ δίνεται με την εντολή ecdf, οπότε πρώτα θα ορίσουμε αυτή την συνάρτηση και μετά θα την προσαρμόσουμε στα δεδομένα μας.

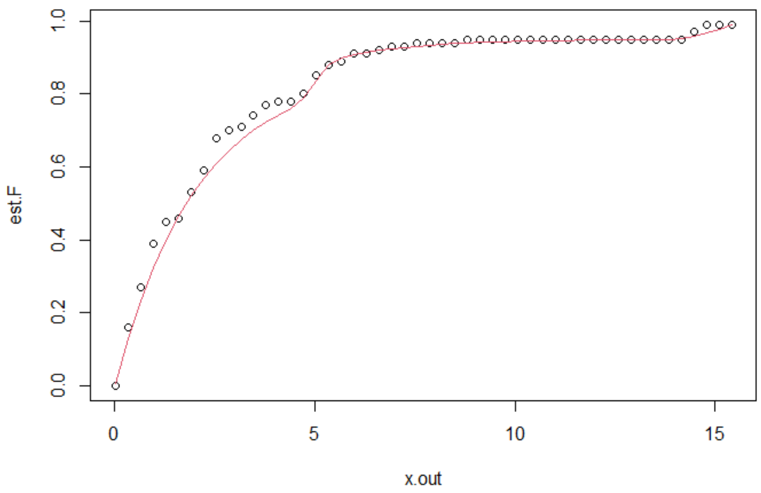
```
ecdf1 <- function (xin,xout) #xin το δείγμα ,xout τα σημεία στα οποία θέλω την τιμή f(x)
{ ΟΤΙΔΙΠΟΤΕ ΕΙΝΑΙ ΜΕΡΟΣ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ ΜΠΑΙΝΕΙ ΣΤΙΣ ΑΓΚΥΛΕΣ
  n <- length (xin)
  nn <- length (xout)
#τα σημεία στα οποία είναι η τιμή της συνάρτησης
  y.ecdf <- vector( length = nn )
  for ( i in 1:nn )
  {
    y.ecdf[i] <- length ( which( xin < xout[i] ) ) / n #τύπος της εμπειρικής ασκ
  }
  y.ecdf #το αποτέλεσμα που θέλω να μου επιστρέφει η συνάρτηση
}
Assign_1_data_1 <- Assign_1_data_1[ ,1] #επικαλούμε τα δεδομένα μου
x.out <- seq ( min ( Assign_1_data_1 ), max ( Assign_1_data_1 ), length = 50 ) #ορίζω από που μέχρι που θα είναι οι τιμές μου
est.F <- ecdf1 ( Assign_1_data_1, x.out ) # υπολογισμός της F καπελάκι (εκτιμήτρια)
plot( x.out , est.F ) # Γραφική παράσταση της εκτιμήτριας.
```



   B)

   Ζωγραφίστε την πραγματική F που είναι μια μίξη κατανομών: κατά 85% μια χ^2 με 2 βαθμούς ελευθερίας, κατά 10% μια $N(5,0.3^2)$ και κατά 5% μια $N(15,0.5^2)$.Ανήκει η F στο παραπάνω διάστημα;

```
true.F <- function( x ) #πραγματική ασκ
{
  0.85*pchisq( x , df = 2 ) + 0.1*pnorm( x, mean = 5, sd = 0.3 ) +
0.05*pnorm( x, mean = 15, sd = 0.5 )
}
lines( x.out , true.F ( x.out ) , col = 2 )
```



Γ) Ελέγξτε σε επίπεδο 5% με το τέστ των Kolmogorov-Smirnov αν τα δεδομένα προέρχονται από μια χ^2 με 2 βαθμούς ελευθερίας. Σε ποιο σημείο χ έχουμε μέγιστη διαφορά μεταξύ της εμπειρικής ασκ και της υποτιθέμενης ασκ.

```
Dn <- ks.test( Assign_1_data_1, "pchisq", df = 2 )
Dn
```

Δ) Ελέγξτε σε επίπεδο 5% με το τέστ των Kolmogorov-Smirnov αν τα δεδομένα προέρχονται από την πραγματική F . Σε ποιο σημείο χ έχουμε μέγιστη διαφορά μεταξύ της εμπειρικής ασκ και της υποτιθέμενης ασκ.

```
Dn1 <- ks.test ( Assign_1_data_1, "true.F" )
Dn1 # εμφανίζει το αποτέλεσμα
```

Ε) εκτιμήστε την F(5)-F(2) είναι κοντά στην πραγματική;

```
ecdf1( Assign_1_data_1 , 5 ) - ecdf1 ( Assign_1_data_1 , 2 ) #υπολογιζουμε την τιμή της εκτιμήτριας
true.F(5) - true.F(2) # Υπολογίζουμε την πραγματική τιμή
```

Συγκρίνουμε.

2. Προσομοιώστε 1000 δείγματα μεγέθους n=100 από μια N(0,1) και ελέγξτε για καθένα ποιο ποσοστό δειγμάτων απορρίπτεται. Χρησιμοποιείστε set.seed(αριθμό των data)

   A) όταν κάνω σε επίπεδο 95% τον έλεγχο K-S για την μηδενική ότι το δείγμα προέρχεται από την N(0,1)

```
set.seed( 100 )
v.pvalue <- vector ( length = 1000 ) # το διάνυσμα που θα αποθηκεύει τις τιμές
for (i in 1:1000)
{
 x.r <- rnorm (15 )
 v.pvalue[i] <- ks.test ( x.r, "pnorm" )$p.value
}
v.pvalue

length  ( which ( v.pvalue  <  0.05 ) ) / 1000
```

   B) Όταν κάνω σε επίπεδο 95% τον έλεγχο K-S για την μηδενική αν το δείγμα προέρχεται από την $\mathrm{N}(\bar{X}, S_n^2)$

```
set.seed( 15 )
v.pvalue <- vector ( length = 1000 )
for (i in 1:1000)
{
 x.r <- rnorm ( 100 )
 v.pvalue[i] <- ks.test ( x.r,"pnorm",mean = mean( x.r ),sd = sd(x.r) )
$p.value #εκτίμηση των παραμέτρων με την μέθοδο της μέγιστης πιθανοφάνειας
}
v.pvalue

length ( which ( v.pvalue < 0.05 ) ) / 1000
```