0.1   Kolmogorov - Smirnov Test

For a single sample ofdata, the Kolmogorov-Smirnov test is used to test whether or not the sample of data is consistent with a specified distribution function. (Not part of this course) When there are two samples of data, it is used to test whether or not these two samples may reasonably be assumed to come from the same distribution. The null and alternative hypotheses are as follows:

H0: The two data sets are from the same distribution
H1: The data sets are not from the same distribution

Consider two sample data sets X and Y that are bothnormally distributed with similar means and variances.

```
> X=rnorm(16,mean=20,sd=5)
> Y=rnorm(18,mean=21,sd=4)
> ks.test(X,Y)


        Two-sample Kolmogorov-Smirnov test

data:  X and Y
D = 0.2153, p-value = 0.7348
alternative hypothesis: two-sided
```

Remark: It doesnt not suffice that both datasets are from the same distribution. They must have the same value for the defining parameters. Consider the case of data sets; X and Z. Both are normally distributed, but with different mean values.

```
> X=rnorm(16,mean=20,sd=5)
> Z=rnorm(16,mean=14,sd=5)
> ks.test(X,Z)


        Two-sample Kolmogorov-Smirnov test

data:  X and Z
D = 0.5625, p-value = 0.0112
alternative hypothesis: two-sided
```