Ex1)

1. In order to perform the required test veridy the assumption of bivariate gaussian via the mcshapiro test whose pvalue is of 0.29 so we can accept the hypothesis of gaussianity.

Test: H0: mu == mu0 vs H1: mu != mu0

With mu0 = (50,50)

At level 5% we reject H0 and so we state that the difference is different from 0.

The pvalue of the test is 0.

1. CR = { m \in R^2 t.c. n \* (x.mean-m)' %\*% x.invcov %\*% (x.mean-m) < cfr.fisher }

* Center

PM2.5 PM10

164.9005 150.4848

* Direction of the axis

[,1] [,2]

[1,] 0.6291343 -0.7772966

[2,] 0.7772966 0.6291343

* Length of the semi-axes

26.295269 8.525918

1. The vector (50,50) is outside the CR indeed we know that we reject H0 iff mu0 does not belong to the CR (equivalently iff the sample mean belongs to the rejection region)
2. T2 simultaneous CI for the 2 components of the mean:

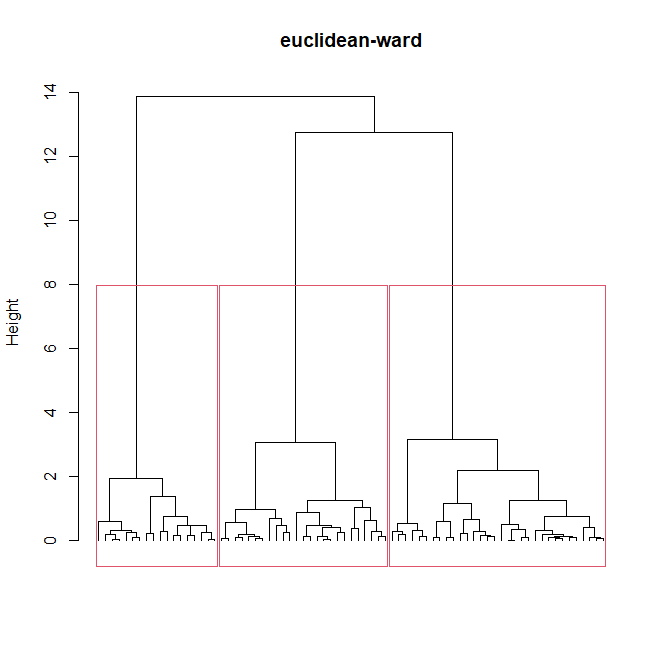
inf center sup

PM2.5: 147.0792 164.9005 182.7218

PM10: 129.3534 150.4848 171.6161

Ex2)

1. Dendogram



I have plotted the division into 3 clusters since from the dendogram this seems the more appropriate and robust choice, moreover it is also coherent with the scatterplot of the data.

1. Verify the hyp:

* Normality in each cluster: we perform a mcshapiro test in each cluster and the pvalue are

0.9252 0.8308 0.3792

So we can assume gaussianity

* Same covariance structure:

Compare both the values qualitatively and visually plotting the matrices -> the assumotion seems to be met

Model:

Model: X.ij = mu + tau.i + eps.ij; eps.ij~N\_p(0,Sigma^2), X.ij, mu, tau.i in R^p

### Test:

### H0: tau.1 = ... = tau.g = (0,0)'

### H1: (H0)^c

### that is

### H0: The membership to a group hasn't any significant effect on the mean

### of X.ij (in any direction of R^2)

### H1: There exists at least one direction in R^2 along which at least two groups

### have some feature significantly different

We obtain a pvalue around the numerical 0 and so we conclude that there is statical evidence to state that the membership to a cluster has an effect on the mean features of the stone flakes.

1. Bonf interavls for the difference of the components of the mean in each cluster:

$group1\_group2

inf12 sup12

Length -0.2216299 0.1474938

Width -3.1823644 -2.4733238

$group1\_group3

inf13 sup13

Length 2.073422 2.4078408

Width -1.193064 -0.5506875

$group2\_group3

inf23 sup23

Length 2.084059 2.471340

Width 1.584009 2.327927

We know that if an interval does not contain the 0 this means that there is statistical difference between the means of that variable in the 2 different groups.

In this specific case we observe that all the intervals do not contain the 0 except for the one related to variable length and clust1 and clust2.

So we state that the difference of the mean of width in each pair of groups is statistically different and also the difference of the mean of the length between group 1 and group 3 and between group2 and group3.

In particular group 1 has a larger length wrt group3 but a lower width.

Group 2 has larger length and width wrt group3.

Group1 has lower width wrt group2 and it does not seem to have a difference wrt the variable length.

Ex3)

1. # Model:

sound = beta0 + beta1\*frequency + beta2\*dummy + beta3\*dummy\*frequency + eps

with eps~ N(0, sigma^2)

where dummy=1 if velocity==’L’

Fitting the model we obtain these estimates for the parameters:

* Beta

(Intercept) frequency

L 123.3592 0.009761214

H 241.9234 0.012276004

We see that both intercept and slope change in the 2 groups due to the presence of the dummy.

* Sigma^2 = 1922.932

1. 1. Perform a test to check if both the parameters related to frequency, so what we called beta1 and beta3 in the above expression, can be assumed = 0

H0: (beta1, beta3) == (0, 0) vs H1: (beta1, beta3) != (0, 0)

The pvalue is almost the numerical 0 so we reject H0

* there is a significant dependence of the mean sound level on the air stream frequency

2. Same test but with the coeff related to the dummy

H0: (beta2, beta3) == (0, 0) vs H1: (beta2, beta3) != (0, 0)

The pvalue is almost the numerical 0 so we reject H0

* there is a significant dependence of the mean sound level on the velocity

3. The test check the significance pf the interaction between the velocity and the frequencies, from the output of the model we can already read the pvalue of the test

H0: beta3== 0 vs H1: beta3 != 0

Which is 0.291 so we assume H0 -> there is not significant impact and we can reduce the model

1. reduced model:

sound = beta0 + beta1\*frequency + beta2\*dummy + eps

with eps~ N(0, sigma^2)

paremters :

* Beta

(Intercept) frequency

L 123.3592 0.012276

H 241.9234 0.012276

Now we can observe that only the interceot vary based on the grouping induced by the velocity

* Sigma^2 = 1928.779

1. CI = [393.8384, 441.4352] CI at level 95%

Ex4)

1. We are considering a model which includes a regressor meaning that we are in the case non-stationary (so the mean is not constant over the space)

For delta we fit a spherical model and the parameters estimated are the following

Sill = 507.243

Range = 1165.763

Nugget = 0

A0 = -29.03012

A1 = 0.02092

1. FATTO CON LM

First fit the linear model for p(s0) in the variable distance

Population = beta0 + beta1\*distance + eps

With eps~ N(0, sigma^2)

So under the hypothesis of residuals centered in 0, homoschedastic and gaussian

The point prediction for p\_s0 gives us 6132.35.

Using this value we find the kriging prediction y\*(s0) = 98.43821

1. È il punto b fatto con spatial

Build a spatial model with response population and regressor distance from the duomo so this is a non stationary model and we assume isotropy. The stationary residual is fitted by a spherical model with nugget

pop.s0 = 6091.035

revenue.s0 = 97.57383

1. Var = 247.2182 which is not fully representative since we are basing this estimate on the prediction of pop.s0 which is itself affected by variability.