LAB 1

Basic commands (scalars, vectors, matrices, and operations)

249 Import/Export of dataframes

Examples of univariate statistical analyses with plots

Visualization of Multivariate Data

Visualization of Categorical Data

3d plots, functions, "for" cycles

Save plots

194 Categorical variables

318 t test

336 simple linear regression

360 command lm

378 test F

LAB 2

Probability density functions, Cumulative distribution functions, Quantiles

Random number generation

QQplots

66 generation random numbers

92 QQplot

211 Shapiro test p-value > 0.05 ok “normal”

LAB 3

Principal Component Analysis

39 variance along direction

165 loadings

209 standardize

279 color factors categorical

297 subset

363 PCA on correlation matrix

409 only significant

429 projection con space generated first k PC

443 Scores

458 p-dimensional geometrical interpretation

LAB 4

Testing for multivariate normality

44 approach 1: linear combination original variables

102 approach 2: squeared mahalanobis distance->chi-squear distribution?

127 chi.sq goodness-of-fit test

143 approach 3: all direction simultaneously min Shapiro-wilk

215-220 rule

332 remove outliers

LAB 5

Box-Cox transformations

320 Tests and confidence regions for the mean of a multivariate Gaussian

60 lambda rule for box-cox

77 univariate Box-Cox transformation

84 trasformation optimal lambda

93 multivariate Box-Cox

132 Bivariate Box-Cox

344 mean of a multivariate Gaussian

373 rejection region

423 Asymptotic test on the mean

515 Bonferroni intervals

589 Confidence region mean 95%

LAB 6

Test for the mean of paired multivariate Gaussian observations

Test for repeated measures

Test for two independent Gaussian populations

57-58 DBOD-DSS

65 Hotelling test

99 Ellipsoidal confidence region 95%

119 Simultaneous T2 intervals

128-187 plot 215-245 superplot

191 look all direction

249 bonferroni più grande più grandi intervalli

406 test for repeated measures

561 test mean indep gaus pop

LAB 7

One-way ANOVA and MANOVA

Two-ways ANOVA and MANOVA

60 verify assumption, normality covariance

79 homogeneity variance barlett test (want high p-value)

84 one-way anova

104 Comparison between groups with bonferroni

159 multiple testing no Bonferroni

224/ 321 One-way manova

408 Two-ways ANOVA

701 general formula for comparison

LAB 8

Linear and Quadratic Discriminant Analysis LDA,QDA

818 Support Vector Machines SVM

14 Example 1 (2 classes, univariate)

38-95 LDA (univariate) 95 do all

158 k-nearest neighbor classifier knn

194 Example 2 (3 classes, bivariate)

222 jittering

270 compute the APER see remark 285

294-310 estimate AER leave-one-out cross-validation see remark 323

353 3d plot

365 Quadratic Discriminand Analysis (QDA)

436 knn-classifier plot

565 fisher discriminant analysis

681 classify new observation

776 plot projections canonical directions (not orth)

824 SVM linear case

871 change cost

890 predict new obs

936 non linear case

LAB 9

19 Hierarchical clustering

434 K-means clustering

Exercises

898 Multidimensional Scaling MDS

34 dissimilarity matrix, chiaro vicini

68 hierarchical clustering hclust()

85 dendrogram

149-164 cophenetic matrices, use coefficents

215 chaining effect (single linkage)

267 ellipsoidal clusters

434-497 K-means method

516 how to choose k

936 most different distances

953 compute the stress

LAB 10

Linear models

18 Multiple linear regression

76 interference on parameters, test Fisher

84 Confidence region

119 Confidence interval

388 dummy variable

816 Multiple linear regression qualitative predictors

LAB 11

Linear models

Problem with collinearity PWC

20/272 Multiple linear regression

51 PWC PCA

155 PWC ridge regression

204 optimal lambda CV

241 PWC lasso regression

258 optimal lambda CV

266 coefficients optimal lambda

308 vif, high => collinearity

310 PCA regression

388 Ridge and Lasso glmnet

475 Variable selection

508 remove NA’s

514 Subset selection

558 ForwardBackward Stepwise Selection

598 choosing k-fold CV

740 Logistic regression

790 Classification and regression trees

884 categorical variable with more than 2 values, dummy variable

LAB 12.1

69 Linear Models homoscedastic and independent errors

133 Linear Models heteroscedastic and independent errors

212 Linear Models heteroscedastic and dependent errors

140 VarIdent() gls

162 VarPower()

199 pearson residuals, standardized

224 variogram increase => correlation decrease

241 Correlation 1: CorCompSym()

282 Correlation 2: AR(1)

308 Correlation 3: general correlation structure

LAB 12.2

53 Linear Mixed Models Homoscedastic Residuals

365 Linear Mixed Models Heteroscedastic Residuals

58 Random intercept, homoscedastic residuals

114 PVRE Percentage Variance explained by Random Effect, intraclass correlation (ICC)

125 prediction

195 random intercept + slope and homoscedastic residuals

196 general D

290 diagonal D

370 same model heteroschedastic

LAB 12.3

95 Add Linear Mixed Effects Models

117 Variance components

222 Prediction

262 LMM Random Intercept & Slope

LAB 13

Exploratory Analysis & Variogram Estimation

67 Exploratory Analysis

82 Variogram Analysis

123 Variogram modeling

176 SSErr sum of squares error

177 weights

208 Spatial Prediction & Kriging

257 Non-stationary Univariate Spatial Prediction (Universal Kriging)

LAB 14.smoothing

60 Regression Splines

86 Fit via LS

135 Approximate pointwise confidence intervals

228 Generalized cross-validation

241 Bias-Variance tradeoff

267 Smoothing Splines

389 Local Polynomial Regression

493 Smoothing For Positive Curves

513 Smoothing For Monotone Curves

635 Extension To Multidimensional Curves

LAB 14.kma

43 without alignment

82 alignment

112 shift and dilation

116 warping function

129 choose number of cluster

LAB 14.FPCA

32 set a high dimensional basis

39 reduced dimensionality

46 compromise between 1 and 2 (32 and 39)

51 estimate of the mean and of the covariance kernel

124 FPCA Functional Principal Component Analysis

147/160/260 plot of the FPCs as perturbation of the mean