0914 code

# Sample R code for reading in data and

# doing basic analyses

# Read in data

wcgs = read.table("wcgs2.dat",header=T)

# Basic descriptives

summary(wcgs)

age height weight sbp

Min. :39.00 Min. :60.00 Min. : 78 Min. : 98.0

1st Qu.:42.00 1st Qu.:68.00 1st Qu.:155 1st Qu.:120.0

Median :45.00 Median :70.00 Median :170 Median :126.0

Mean :46.28 Mean :69.78 Mean :170 Mean :128.6

3rd Qu.:50.00 3rd Qu.:72.00 3rd Qu.:182 3rd Qu.:136.0

Max. :59.00 Max. :78.00 Max. :320 Max. :230.0

smoke

Min. :0.0000

1st Qu.:0.0000

Median :0.0000

Mean :0.4762

3rd Qu.:1.0000

Max. :1.0000

# variance-covariance matrix

> var(wcgs)

age height weight sbp smoke

age 30.515074 -1.332266318 -4.009296 13.84165155 0.013130996

height -1.332266 6.394290367 28.429290 0.70238689 -0.004493913

weight -4.009296 28.429289696 445.031158 80.76637986 -1.271954260

sbp 13.841652 0.702386892 80.766380 228.54577768 0.019810827

smoke 0.013131 -0.004493913 -1.271954 0.01981083 0.249513654

> cor(wcgs)

age height weight sbp smoke

age 1.000000000 -0.095375682 -0.03440454 0.165746397 0.004758754

height -0.095375682 1.000000000 0.53293547 0.018373573 -0.003557799

weight -0.034404537 0.532935466 1.00000000 0.253249623 -0.120706075

sbp 0.165746397 0.018373573 0.25324962 1.000000000 0.002623426

smoke 0.004758754 -0.003557799 -0.12070607 0.002623426 1.000000000

pairs(wcgs)

pairs(wcgs,pch="o")

# Bivariate scatterplots

scatter.smooth(wcgs$age,wcgs$height)

install.packages("lattice")

library("lattice")

# Trivariate scatterplots

cloud(sbp ~ age \* height, data = wcgs)

cloud(sbp ~ age \* height | smoke, data = wcgs)

# qqplot: assessing normality

qqnorm(wcgs$sbp)

qqline(wcgs$sbp,col=2)

0912

# R code for computing correlation from a variance-

# covariance matrix

a = matrix(rnorm(100),50,2)

# Sample covariance matrix

cmat = var(a)

# Construct D^{-1/2} from the notes

dmat.neghalf = diag(diag(cmat)^(-1/2),2,2)

# Sample correlation

cor.mat = dmat.neghalf %\*% cmat %\*% dmat.neghalf

# Compare to correlation matrix

cor.mat2 = cor(a)