Computational Statistics and Data Analysis
Sheet No. 6

David Bubeck, Patrick Nisblè

May 31, 2017

1 Correlated Variables

$$\bar{V}_1 = 6.991867 \tag{1}$$
$$\bar{V}_2 = 9.009960 \tag{2}$$

We calculated the following two matrices:

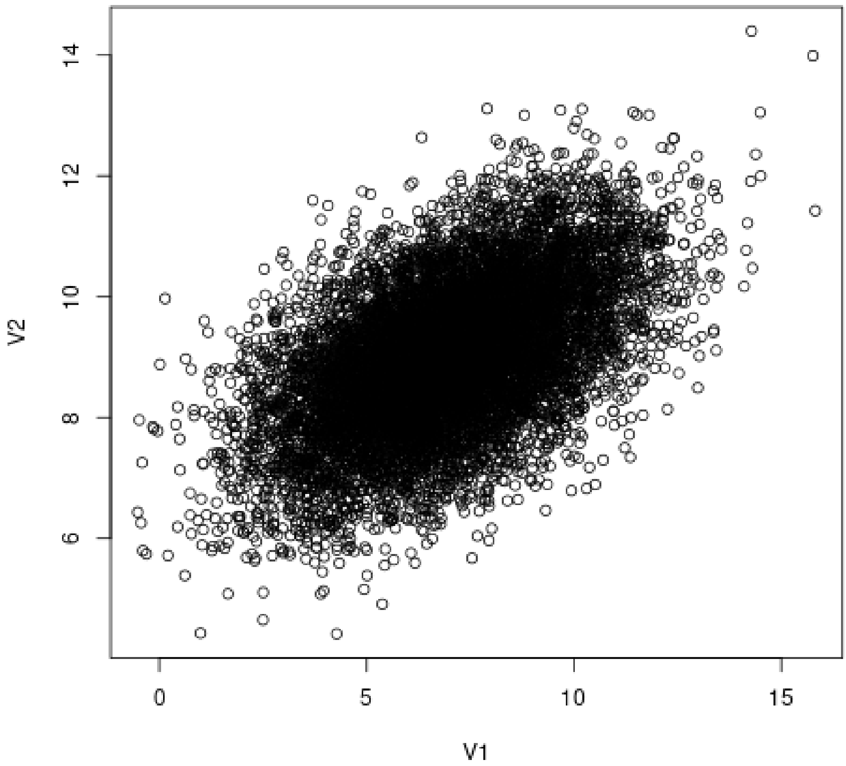
$$\text{covariance matrix } C_{ij} = \begin{pmatrix} 5.045109 & 1.508497 \\ 1.508497 & 1.492933 \end{pmatrix} \tag{3}$$
$$\text{correlation matrix } C_{ij} = \begin{pmatrix} 1.0000000 & 0.5496535 \\ 0.5496535 & 1.0000000 \end{pmatrix} \tag{4}$$

using following code:

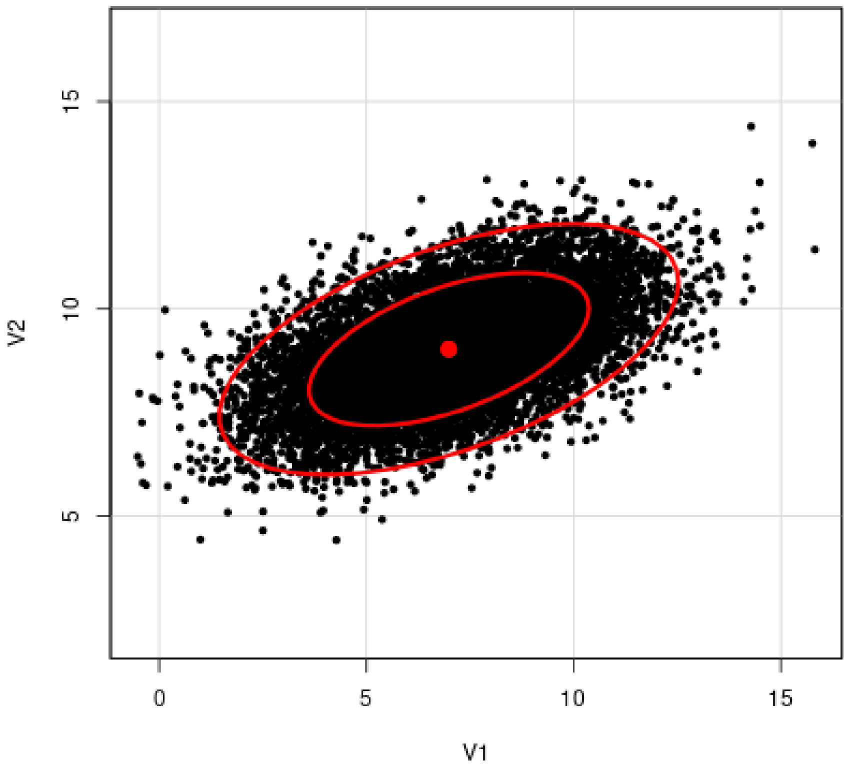
```r

library(MASS)
library(mnormt)
library(car)

#import data
data <- read.table("bivariate-measurements.txt", header = TRUE, sep = " ")

n_row = dim(data)[1]
n_col = dim(data)[2]

#calculating mean vector
mean <- matrix(0, nrow = 2, ncol = 1)
for(col in 1:n_col) {
    for(row in 1:n_row) {
        mean[col] = mean[col] + data[row, col]
    }
    mean[col] = mean[col] / n_row
}

#calculating difference matrix and variance / deviation
diffMat <- matrix(0, nrow = n_row, ncol = n_col)
var <- matrix(0, nrow = n_col, ncol = 1)
dev <- matrix(0, nrow = n_col, ncol = 1)
for(col in 1:n_col) {
    for(row in 1:n_row) {
        diffMat[row, col] = data[row, col] - mean[col]
        #calculating variance for corrolation matrix
        var[col] = var[col] + (data[row, col] - mean[col])^2
    }
    #calculating deviation
    dev[col] = sqrt(1/(n_row - 1) * var[col])
}

#calculating covariance matrix
covMat <- (n_row - 1)^(-1) * t(diffMat) %*% diffMat
covMat

#calculating corrolation matrix
corMat <- matrix(0, nrow = dim(covMat)[1], ncol = dim(covMat)[2])
for(i in 1:dim(covMat)[1]) {
    for(j in 1:dim(covMat)[2]) {
        corMat[i, j] = covMat[i, j] / (dev[i]*dev[j])
    }
}
corMat
```

```r
dataplot = mvrnorm(n_row, mu = mean, Sigma = covMat)
#plotting the data
png('dataPlot.png')
plot(dataplot, xlab = "V1", ylab = "V2")

#plotting ellipses with confidence regions
png('ellipsesPlot.png')
dataEllipse(as.matrix(dataplot), levels = c(0.6827, 0.9545),
                lwd = 3, asp = 1, xlab = "V1", ylab = "V2",
                pch = 20)
```

resulting in:



V2
V1
0
5
10
15
6
8
10
12
14

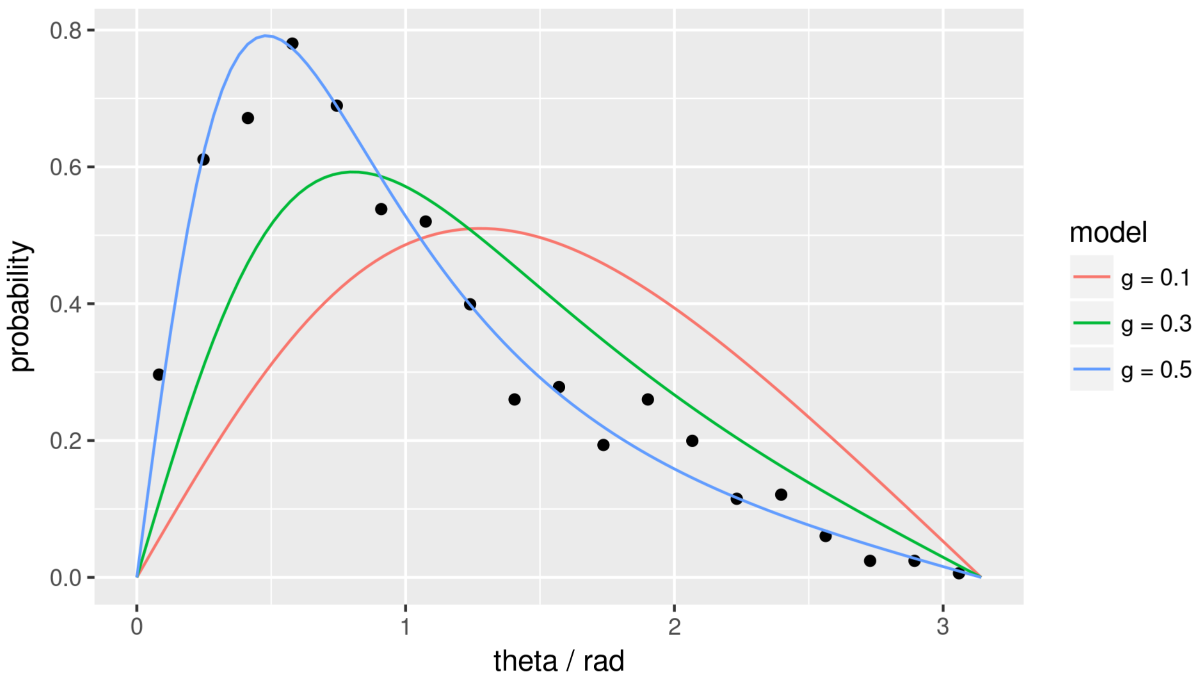


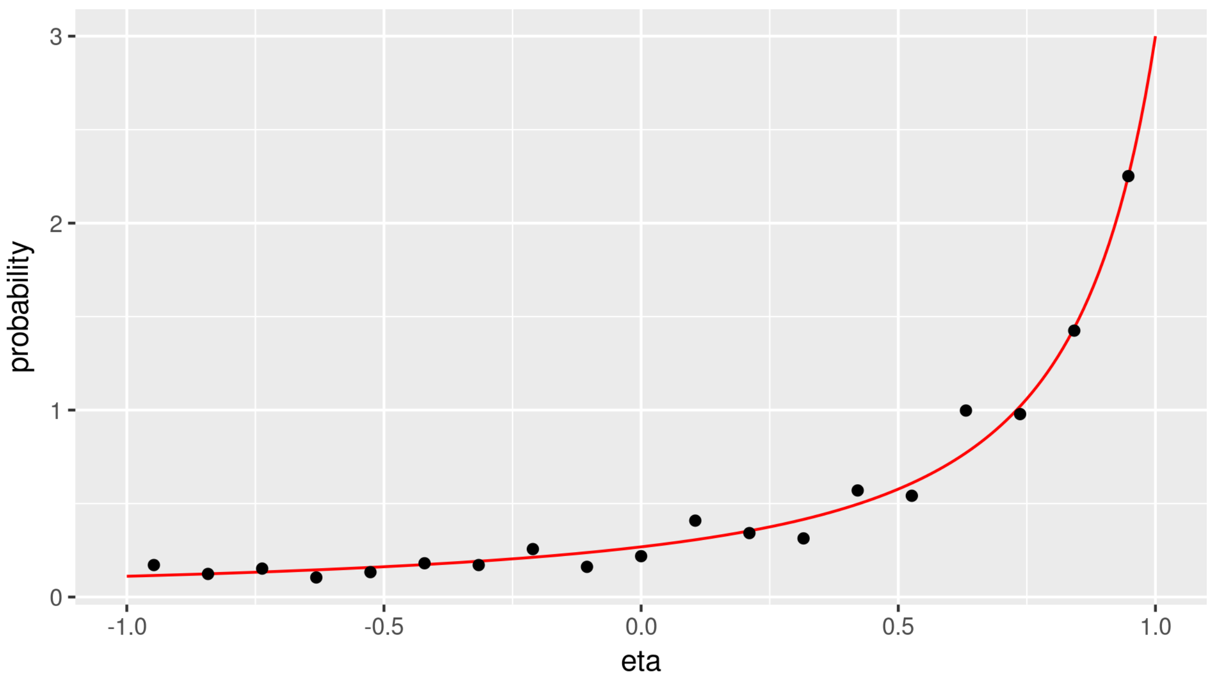
V2
V1
0
5
10
15
5
10
15

2 Non-uniform distributions

using the following code we plotted $p(\theta, g)$ und $p(\eta, g)$, as well as performed sampling with g=0.5

```r
ntests = 1000

# part 1
scatter_prb <- function(theta, g){
    1/(4*pi)*(1-g^2)/(1+g^2-2*g*cos(theta))^(3/2)
}

png('6-2_scatter_prob.png')
xdata = seq(0, pi, length.out=100)
#plot(0,type='n', xlim=c(0,pi), xlab="theta / pi", ylab="p(theta)", xaxt='n')

data = sample(xdata,ntests,rep=TRUE,prob=2*pi*sin(xdata)*scatter_prb(xdata, .5))
h = hist(data, breaks=seq(0,pi, length.out=20))
plot(h$mids, h$density, xaxt='n', xlab="theta / pi", ylab="p(theta)", pch=20)

for (g in seq(.1,.5,.2)) {
    lines(xdata, 2*pi*sin(xdata)*scatter_prb(xdata, g), col=1+g*10)
}
axis(1, at=seq(0,pi,pi/4),labels=seq(0, 1, 1/4), las=2)

#mean and variance
paste('theta_mean = ', mean(data))
paste('theta_var  = ', var(data))

# part 2
scatter_eta <- function(eta, g){
    scatter_prb(acos(eta),g)
}
xdata = seq(-1,1,.01)

png('6-2_scatter_eta.png')
#plot(0,type='n', ylim=c(0,.8*ntests), xlim=c(-1,1), xlab="eta", ylab="p(eta)")
data = sample(xdata,ntests,rep=TRUE,prob=scatter_eta(xdata,.5))
h = hist(data, breaks=seq(-1,1, length.out=20))
plot(h$mids, h$density/4, xlim=c(-1,1), xlab="eta", ylab="p(eta)", pch=20)

lines(xdata, scatter_eta(xdata, .5), col=2)
```


probability
theta / rad
model
g = 0.1
g = 0.3
g = 0.5
0.0
0.2
0.4
0.6
0.8
0
1
2
3


probability
eta
0
1
2
3
-1.0
-0.5
0.0
0.5
1.0

3 Higher order moments of multivariate Gaussian

$$E[x_1 \cdot x_2 \cdot \ldots x_{2n}] = \sum_{1}^{\binom{2n}{2}} \left(\prod_{1}^{n} E[x_i x_j] \right) \tag{5}$$

3.1 a)

$$\begin{aligned} E[x_1^4 x_2^2] =& \binom{3}{1}_{perm} 2^0_{rot} E[x_1^2]^2 E[x_2^2] && (6) \\ &+ \binom{3}{1}_{perm} 2^2_{rot} E[x_1^2] E[x_1 x_2]^2 && (7) \\ =& 3\sigma_1^4 \sigma_2^2 + 12(\varrho_{12}\sigma_1\sigma_2)^2 \sigma_1^2 && (8) \\ & && (9) \end{aligned}$$

3.2 b)

$$\begin{aligned} E[x_1^3 x_2] =& \binom{4}{2} E[x_1^2] E[x_1 x_2] && (10) \\ =& 6\sigma_1^2 (\varrho_{12}\sigma_1\sigma_2) && (11) \end{aligned}$$