R-code for Chapter 1: Multivariate distributions and copulas

Claudia Czado

01 March, 2019

Contents

Required R packages **1**

Section 1.1 Univariate distributions **2**
Figure 1.1: Univariate densities: normal and Student t distribution with nu=3 2

Section 1.2 Multivariate distributions **3**
Figure 1.2: Histograms: (top row) of a standard normal random sample of size 100 (left) and 500 (right) and their associated PIT values using the empirical distribution function (middle row) and the standard normal distribution function (bottom row). 3
Figure 1.3: Bivariate densities and contour lines: left: bivariate normal with zero means, unit variances and rho = .8, middle: contour lines for bivariate normal with rho = :8 (solid lines) and bivariate standard Student t with nu = 3; rho = .8(dotted lines) and right: bivariate standard Student t with nu = 3 and rho = .8. 5

Section 1.3: Features of multivariate data **7**
Ex 1.9: Chemical components of wines 7
Figure 1.4: WINE3: top row: marginal histograms left: acf, middle: acv, right: acc; middle row: pairs plots left: acf versus acv, middle: acf versus acc, right: acv versus acc. bottom row: normalized pairs plots left: acf versus acv, middle: acf versus acc, right: acv versus acc. . . . 8
Figure 1.5: WINE3: top row: marginal histograms of the copula data (left: acf, middle: acv,right: acc), bottom row: pairs plots of the copula data (left: acf versus acv,middle: acf versus acc, right: acv versus acc). 9

Section 1.5: Elliptical copulas **10**
Gauss and Student t copulas objects with common Kendall's tau=.7 (rho=.89) 10
Figure 1.6: Bivariate copula densities: left: GaussGauss copula density surface with rho=.89 . . . 10

Section 1.8: Meta distributions **13**
Figure 1.7: Scatter plots of a bivariate sample (n = 200) from Gaussian meta distributions: left: N (0, 1)/exponential with rate lambda = 5 margins, middle: uniform[0, 1]/uniform[0, 1] margins,right: N (0, 1)/N (0, 1) margins. The common copula is a bivariate Gauss copula with rho = .88. 13

Section 1.10: Exercises **14**

Figure 1.8: WINE7: Scatter plots of the seven dimensional red wine data. **14**

Required R packages

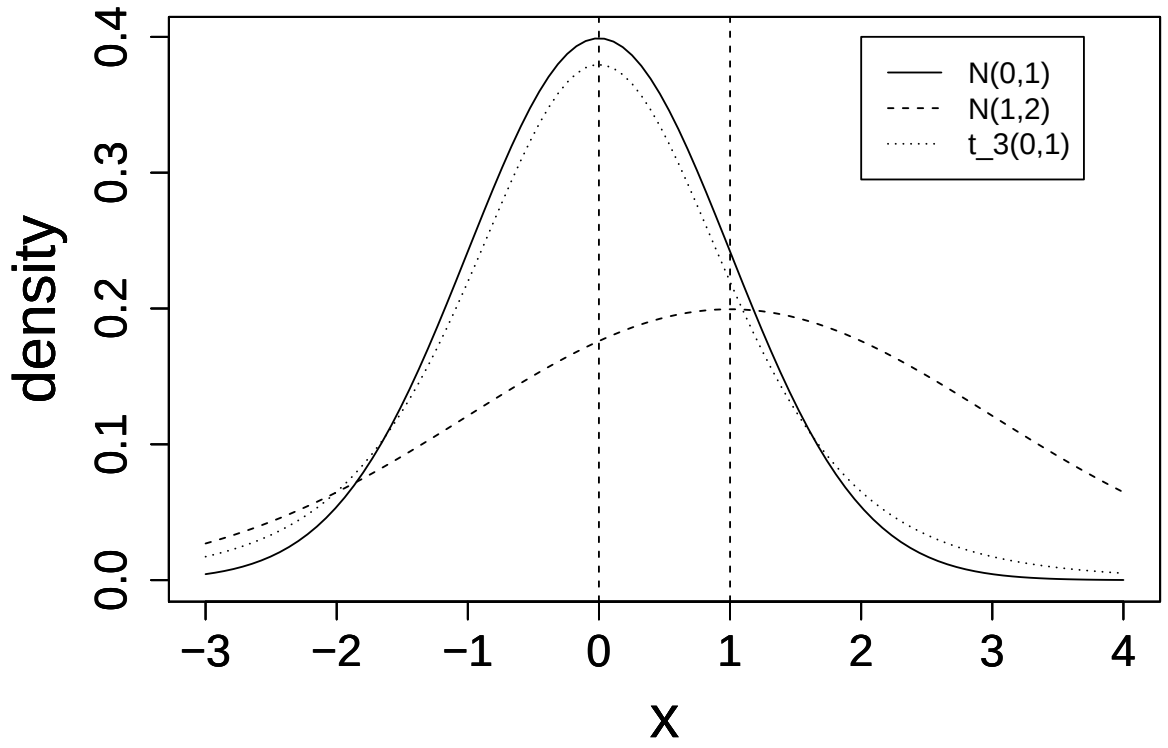
- VineCopula
- rafalib
- copula

- mvtnorm
- lattice

Section 1.1 Univariate distributions

Figure 1.1: Univariate densities: normal and Student t distribution with nu=3

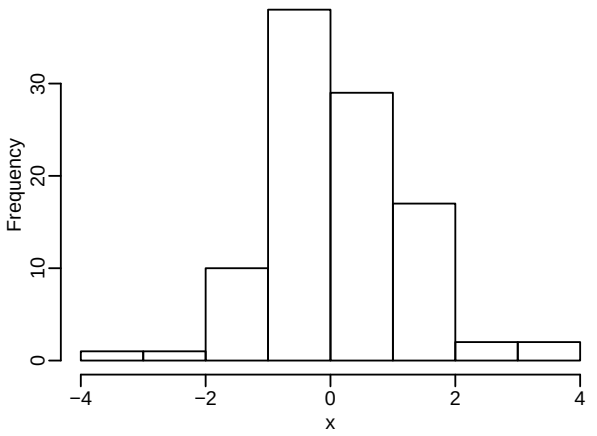
```
x<-seq(-3,4,length=100)
n01.pdf<-dnorm(x)
n12.pdf<-dnorm(x,mean=1,sd=2)
t3.pdf<-dt(x,df=5)
ry=range(n01.pdf,n12.pdf,t3.pdf)
bigpar(1,1)
plot(x,n01.pdf,type="l",ylab="density",lty=1,ylim=ry)
par(new=TRUE)
plot(x,n12.pdf,type="l",ylab="density",lty=2,ylim=ry)
par(new=TRUE)
plot(x,t3.pdf,type="l",ylab="density",lty=3,ylim=ry)
abline(v=0,lty=2)
abline(v=1,lty=2)
legend(2,.4,legend = c("N(0,1)","N(1,2)","t_3(0,1)"),lty=1:3)
```

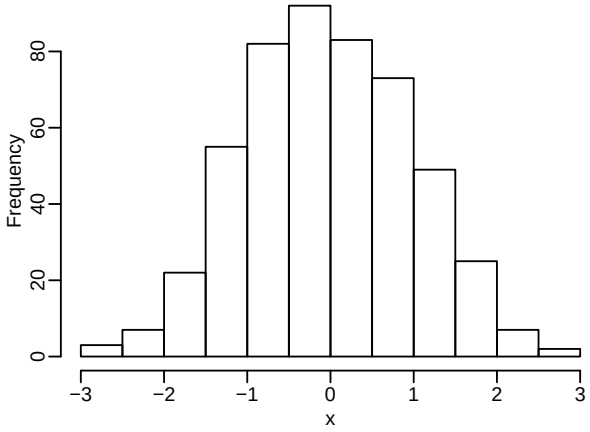


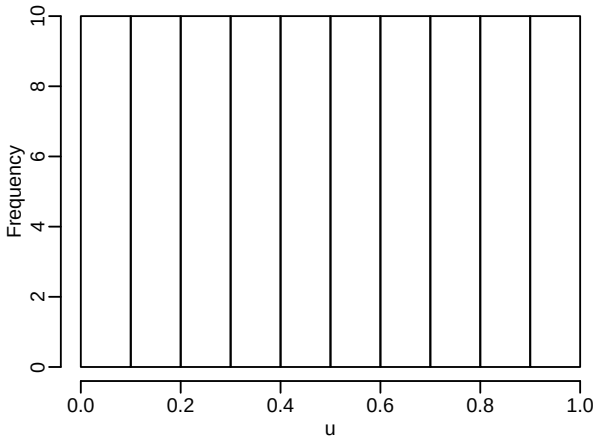
Section 1.2 Multivariate distributions

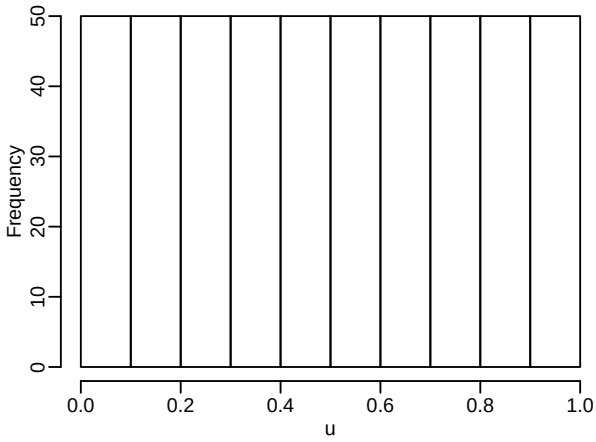
Figure 1.2: Histograms: (top row) of a standard normal random sample of size 100 (left) and 500 (right) and their associated PIT values using the empirical distribution function (middle row) and the standard normal distribution function (bottom row).

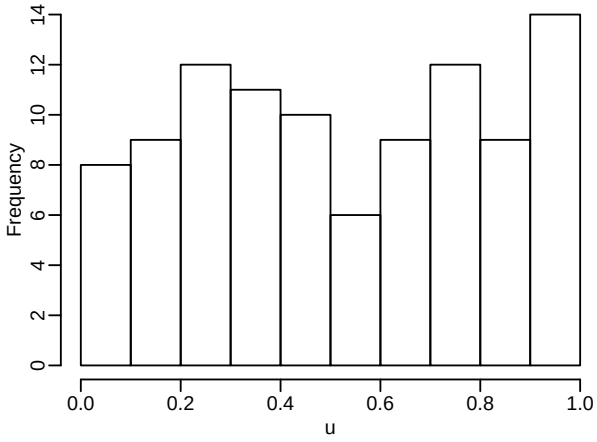
```
rnorm100<-rnorm(100)
rnorm500<-rnorm(500)
pit100<-pnorm(rnorm100)
pit500<-pnorm(rnorm500)
u100<-rank(rnorm100)/101
u500<-rank(rnorm500)/501
mypar(3,2)
hist(rnorm100,main="",xlab="x")
hist(rnorm500,main="",xlab="x")
hist(u100,main="",xlab="u")
hist(u500,main="",xlab="u")
hist(pit100,main="",xlab="u")
hist(pit500,main="",xlab="u")
```


Frequency
0
10
20
30
−4
−2
0
2
4
x


Frequency
0
20
40
60
80
−3
−2
−1
0
1
2
3
x


Frequency
0
2
4
6
8
10
0.0
0.2
0.4
0.6
0.8
1.0
u


Frequency
0
10
20
30
40
50
0.0
0.2
0.4
0.6
0.8
1.0
u


Frequency
0
2
4
6
8
10
12
14
0.0
0.2
0.4
0.6
0.8
1.0
u

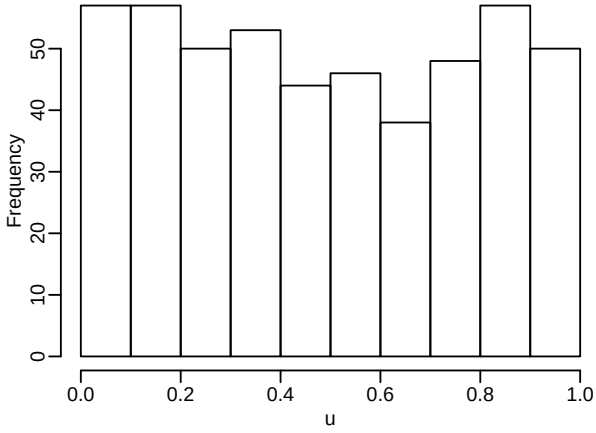

Frequency
0
10
20
30
40
50
0.0
0.2
0.4
0.6
0.8
1.0
u

Figure 1.3: Bivariate densities and contour lines: left: bivariate normal with zero means, unit variances and rho = .8, middle: contour lines for bivariate normal with rho = :8 (solid lines) and bivariate standard Student t with nu = 3; rho = .8(dotted lines) and right: bivariate standard Student t with nu = 3 and rho = .8.

Figure 1.3 (left panel)

```
g <- expand.grid(x = seq(-3,3,length=40), y = seq(-3,3,length=40))
g$z<-dmvnorm(cbind(g$x,g$y),sigma=matrix(c(1,0.8,0.8,1),nrow=2,ncol=2))
trellis.par.set("axis.line",list(col=NA,lty=1,lwd=1))
wireframe(z ~ x * y, data = g,scales = list(arrows = FALSE),xlab="x1",drape=FALSE,
perspective = FALSE, colorkey = FALSE,ylab="x2",zlab="x3")
```

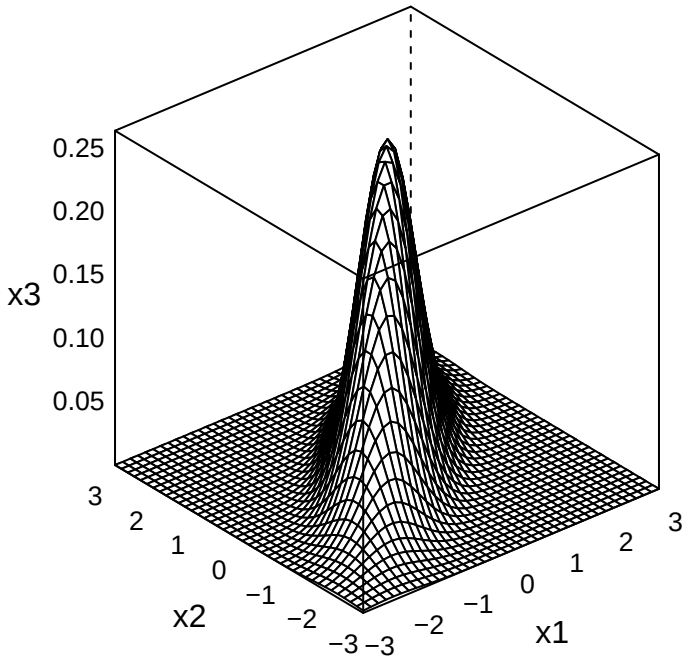


Figure 1.3 (right panel)

```
gt<- expand.grid(x = seq(-3,3,length=40), y = seq(-3,3,length=40))
df=3
Sigma=matrix(c(1,.8,.8,1),nrow=2,ncol=2)
S=Sigma*(df-2)/df
```

```
S=Sigma
gt$z<-dmvt(cbind(gt$x,gt$y),df=df,sigma=S,log=FALSE)
#head(gt$z)
trellis.par.set("axis.line",list(col=NA,lty=1,lwd=1))
wireframe(z ~ x * y, data = gt,scales = list(arrows = FALSE),xlab="x1",drape=FALSE,
perspective = FALSE, colorkey = FALSE,ylab="x2",zlab="x3")
```

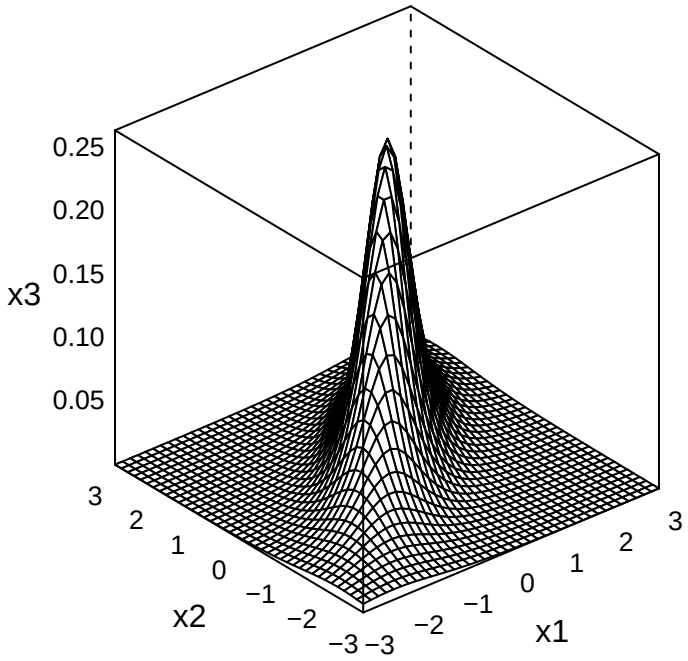
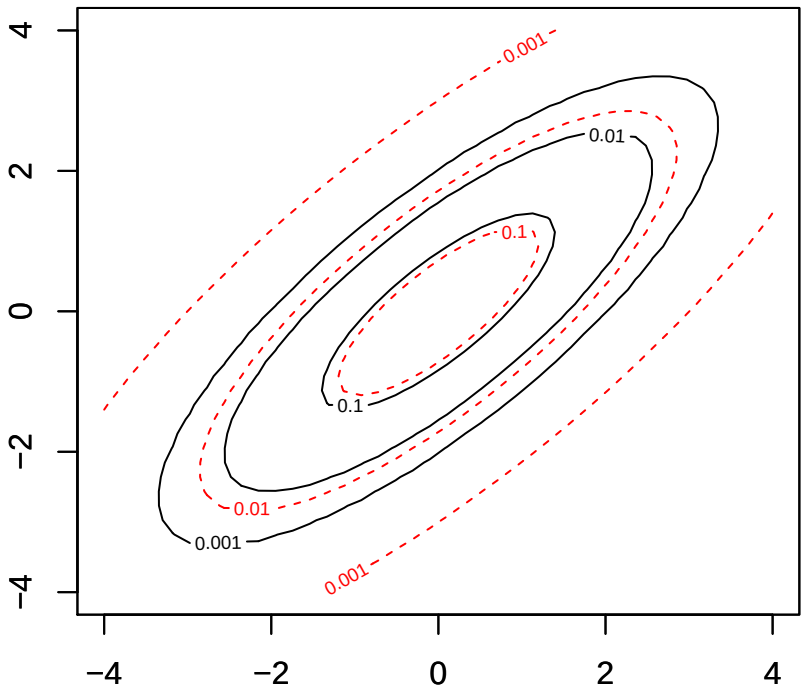


Figure 1.3 (middle panel)

```
x=seq(-4,4,length=40)
y=x
df=3
Sigma=matrix(c(1,.8,.8,1),nrow=2,ncol=2)
S=Sigma*(df-2)/df
S=Sigma
dmt1<-function(x,y){dmvt(cbind(x,y),df=df,sigma=S,log=FALSE)}
dmnorm1<-function(x,y){dmvnorm(cbind(x,y),sigma=Sigma)}
z=outer(x,y,dmnorm1)
zt=outer(x,y,dmt1)
contour(x, y, z,levels=c(0.001,0.01,.1))
par(new=T)
```

```
contour(x, y, zt,levels=c(0.001,.01,.1),lty=2,col=2)
```



Section 1.3: Features of multivariate data

Ex 1.9: Chemical components of wines

Read in data and set column names

```
#reddata<-read.csv(file="C:/Users/cczado/LRZ Sync+Share/vineweb2019/winequality-red.csv",sep=";")
#reddata<-read.csv(file="/home/cczado/linux/Desktop/LRZ Sync+Share/vineweb2019/winequality-red.csv",sep
reddata<-read.csv(file="winequality-red.csv",sep=";")
n<-length(reddata[,1])
colnames(reddata)<-c("acf","acv","acc","sugar","clor","sf","st","den","ph","sp","alc","quality")
```

Transform original data to copula data (udata) using ranks and then to marginal normalized data (zdata)

```
udata<-reddata
zdata<-reddata
```

```
for(i in 1:12){
udata[,i]<-rank(reddata[,i])/(n+1)
zdata[,i]<-qnorm(udata[,i])
}
```

Figure 1.4: WINE3: top row: marginal histograms left: acf, middle: acv, right: acc; middle row: pairs plots left: acf versus acv, middle: acf versus acc, right: acv versus acc. bottom row: normalized pairs plots left: acf versus acv, middle: acf versus acc, right: acv versus acc.

```
bigpar(3,3)
hist(reddata[,1],main="",xlab="acf")
hist(reddata[,2],main="",xlab="acv")
hist(reddata[,3],main="",xlab="acc")
plot(reddata[,1],reddata[,2],pch=".", cex=3,xlab="acf",ylab="acv",
main="")
plot(reddata[,1],reddata[,3],pch=".", cex=3,xlab="acf"
,ylab="acc",main="")
plot(reddata[,2],reddata[,3],pch=".", cex=3,xlab="acv",
ylab="acc", main="")
plot(zdata[,1],zdata[,2],pch=".", cex=3,xlab="acf(normalized)",ylab="acv(normalized)",
main="")
plot(zdata[,1],zdata[,3],pch=".", cex=3,xlab="acf(normalized)"
,ylab="acc(normalized)",main="")
plot(zdata[,2],zdata[,3],pch=".", cex=3,xlab="acv(normalized)",
ylab="acc(normalized)", main="")
```

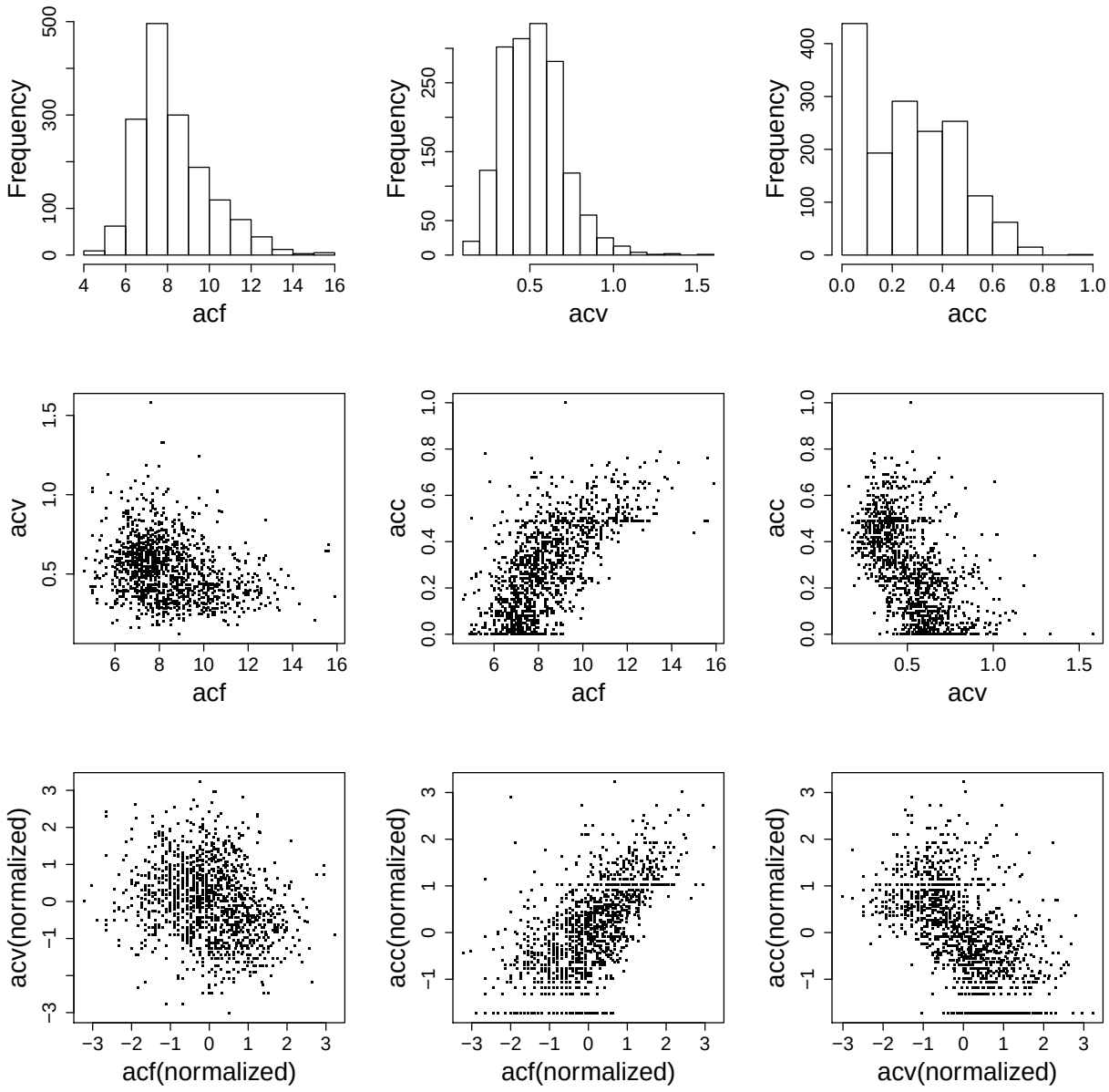
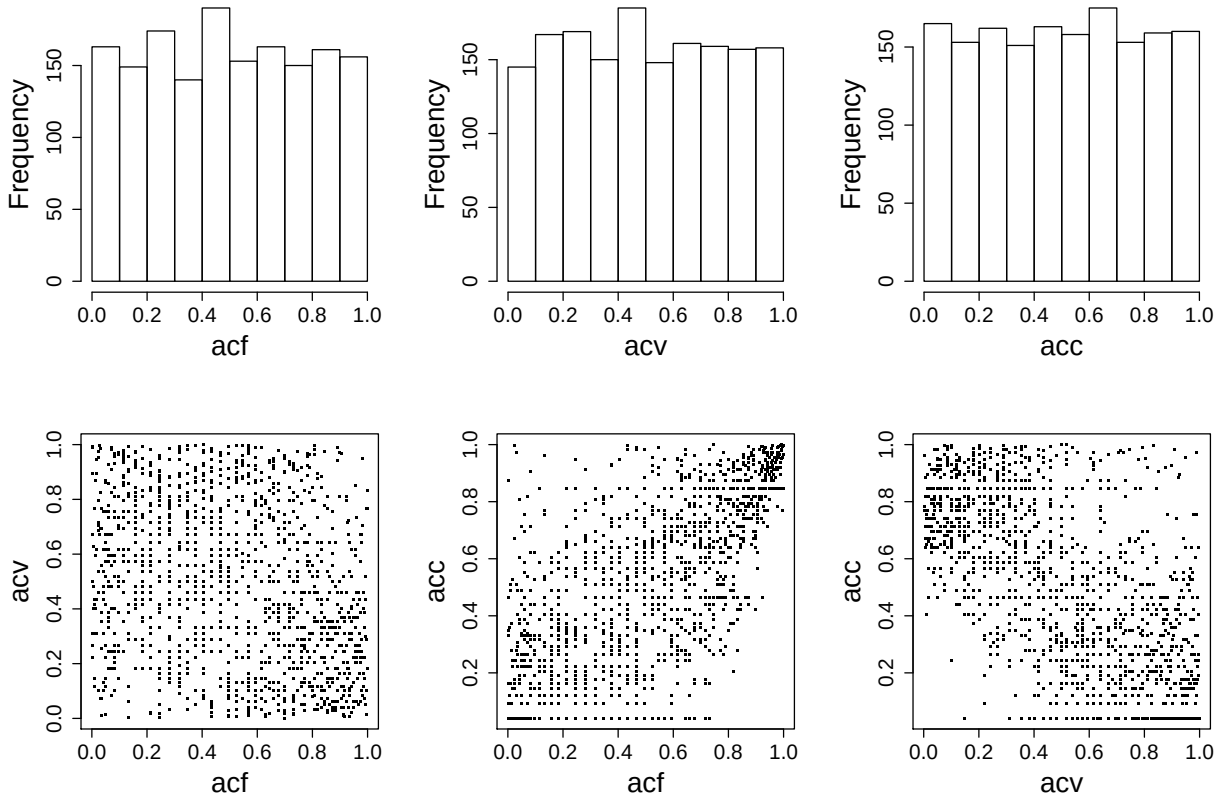


Figure 1.5: WINE3: top row: marginal histograms of the copula data (left: acf, middle: acv,right: acc), bottom row: pairs plots of the copula data (left: acf versus acv,middle: acf versus acc, right: acv versus acc).

```
bigpar(2,3)
hist(udata[,1],main="",xlab="acf")
hist(udata[,2],main="",xlab="acv")
hist(udata[,3],main="",xlab="acc")
plot(udata[,1],udata[,2],pch=".", cex=3,xlab="acf",ylab="acv",
main="")
plot(udata[,1],udata[,3],pch=".", cex=3,xlab="acf"
,ylab="acc",main="")
```

```
plot(udata[,2],udata[,3],pch=".", cex=3,xlab="acv",
ylab="acc", main="")
```



Section 1.5: Elliptical copulas

Gauss and Student t copulas objects with common Kendall's tau=.7 (rho=.89)

```
par.gauss<-BiCopTau2Par(1, .7, check.taus = TRUE)
par.st3<-BiCopTau2Par(1, .7, check.taus = TRUE)
par.st8<-BiCopTau2Par(1, .7, check.taus = TRUE)
obj.gauss <- BiCop(family = 1, par = par.gauss)
obj.st3 <- BiCop(family = 2, par = par.st3,par2=3)
obj.st8 <- BiCop(family = 2, par = par.st8,par2=8)
```

Figure 1.6: Bivariate copula densities: left: GaussGauss copula density surface with rho=.89

```
plot(obj.gauss,zlim=c(0,12))
```

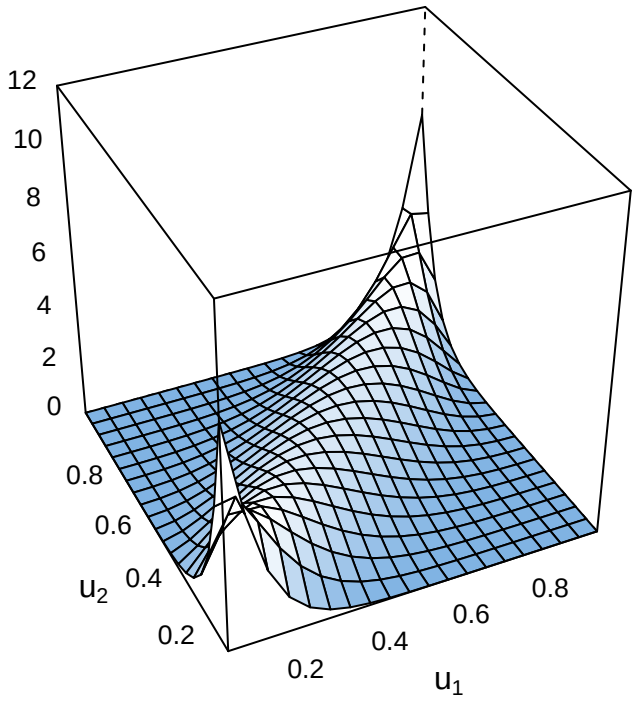


Figure 1.6: Bivariate copula densities: middle: Student t with df=3 and rho=.89

```
plot(obj.st3,zlim=c(0,12))
```

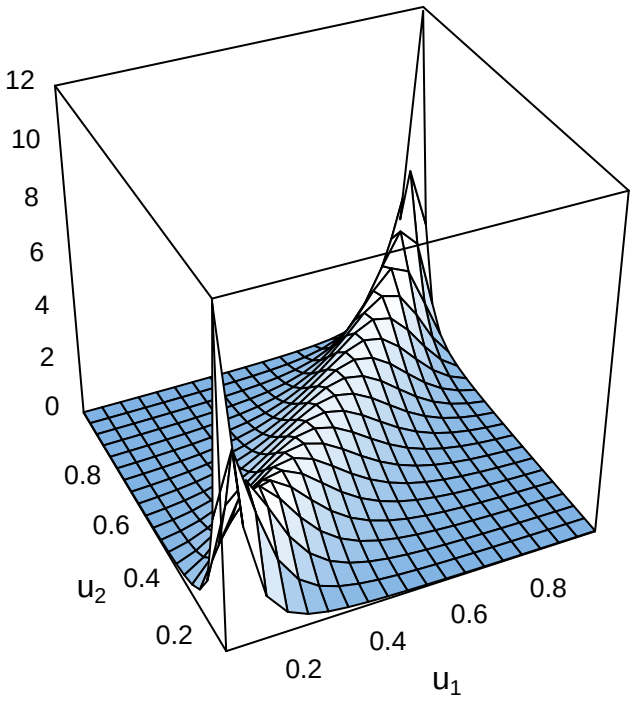
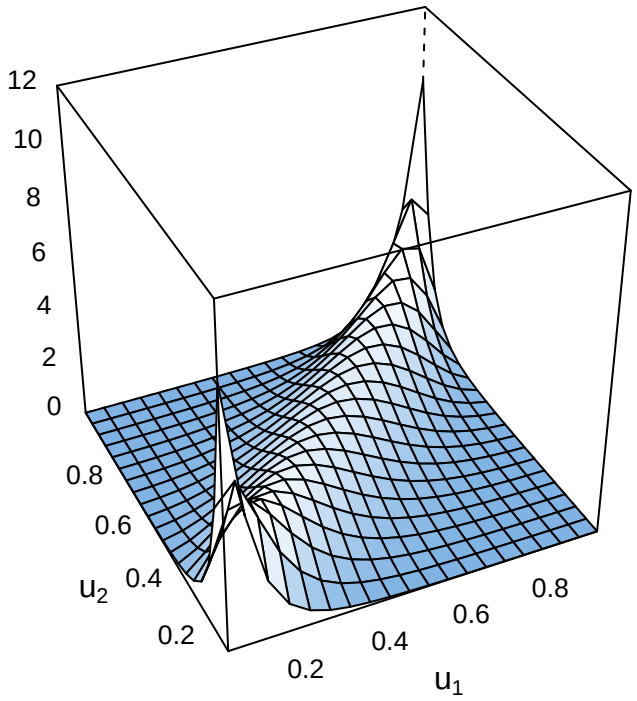


Figure 1.6: Bivariate copula densities: right: Student t with df=8 and rho=.89

```
plot(obj.st8,zlim=c(0,12))
```



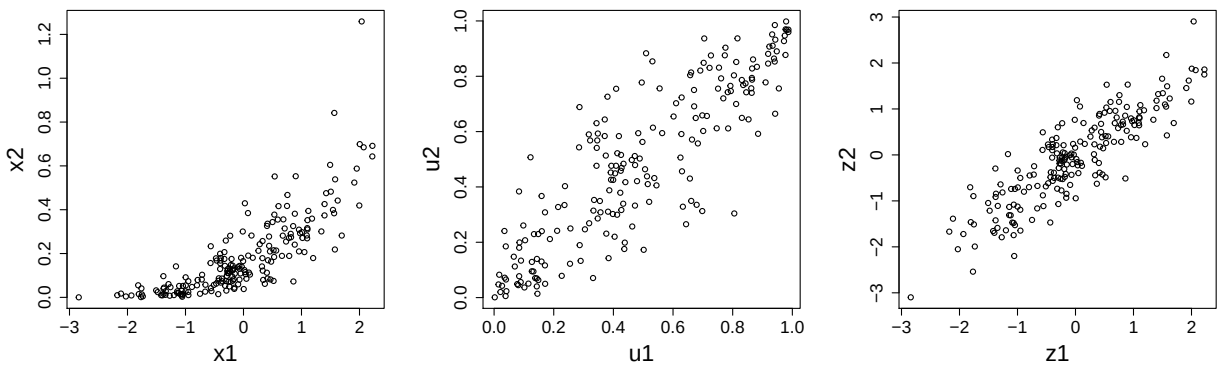
Section 1.8: Meta distributions

```
umeta<-BiCopSim(N=200,family=1,par=par.gauss)
x1<-qnorm(umeta[,1])
u1<-umeta[,1]
u2<-umeta[,2]
x2<-qexp(umeta[,2],rate=5)
z1<-x1
z2<-qnorm(umeta[,2])
```

Figure 1.7: Scatter plots of a bivariate sample (n = 200) from Gaussian meta distributions: left: N (0, 1)/exponential with rate lambda = 5 margins, middle: uniform[0, 1]/uniform[0, 1] margins,right: N (0, 1)/N (0, 1) margins. The common copula is a bivariate Gauss copula with rho = .88.

```
bigpar(1,3)
plot(x1,x2)
plot(u1,u2)
```

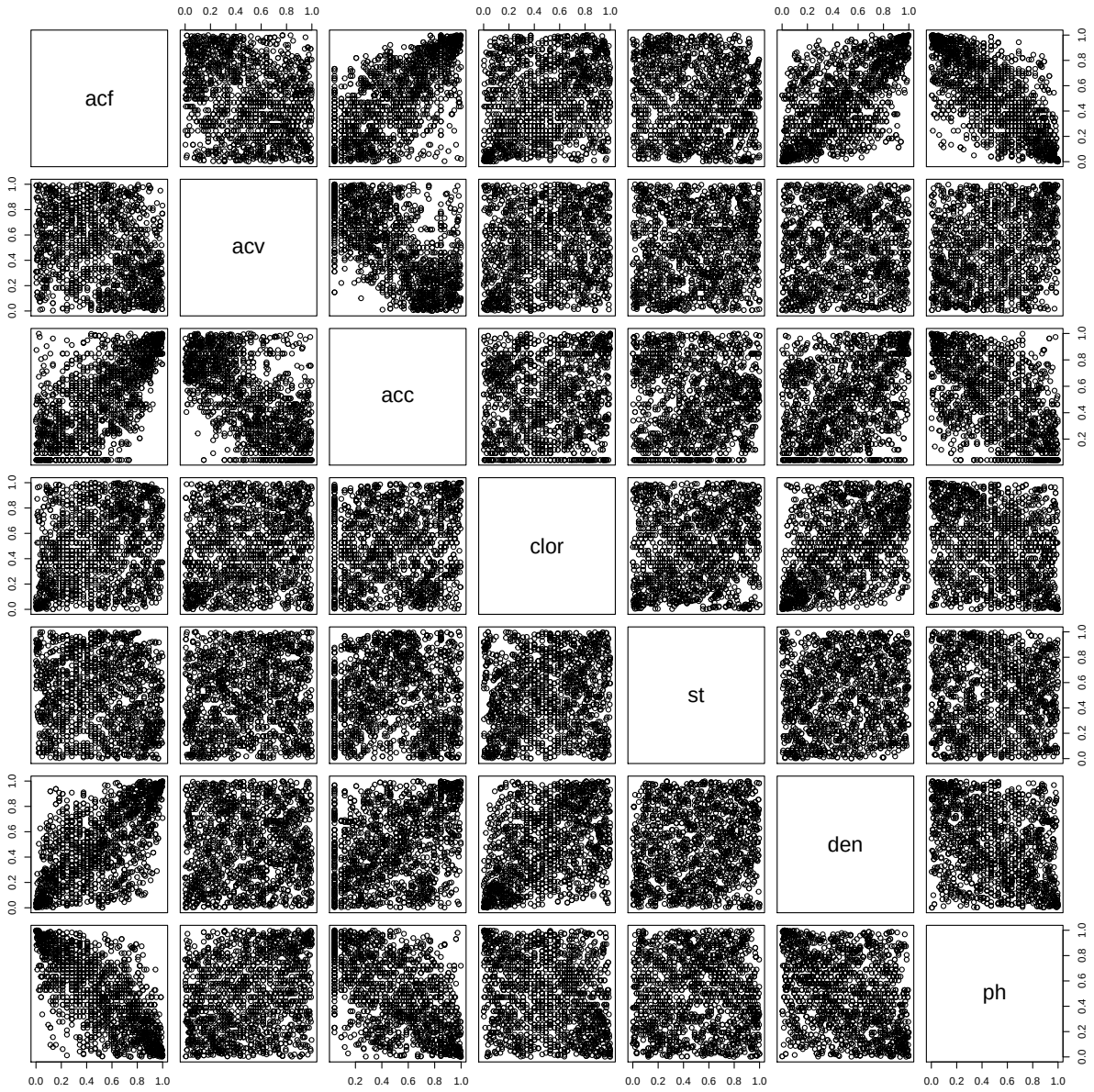
```
plot(z1,z2)
```



Section 1.10: Exercises

Figure 1.8: WINE7: Scatter plots of the seven dimensional red wine data.

```
udata7<-udata[,c(1,2,3,5,7,8,9)]
pairs(udata7)
```


acf
acv
acc
clor
st
den
ph
0.0 0.2 0.4 0.6 0.8 1.0