Multivariate Analysis for the Behavioral Sciences, Second Edition (Chapman and Hall/CRC, 2019)

Examples of Chapter 12: Multivariate Data and Multivariate Analysis

Kimmo Vehkalahti and Brian S. Everitt

28 November 2018

Contents

Examples **2**

Table 12.1: Chest, Waist, and Hip Measurements of 20 Individuals **2**
Figure 12.1 . . . 3
Figure 12.2 . . . 4
Figure 12.3 . . . 5
Figure 12.4 . . . 7
Figure 12.5 . . . 8

Table 12.2: Data on Paint Sprayers **9**
Figure 12.6 . . . 10
Figure 12.7 . . . 11
Figure 12.8 . . . 12

Examples

Table 12.1: Chest, Waist, and Hip Measurements of 20 Individuals

```
body <- structure(list(
  Chest = c(34, 37, 38, 36, 38, 43, 40, 38, 40, 41, 36, 36, 34, 33, 36, 37, 34, 36, 38, 35),
  Waist = c(30, 32, 30, 33, 29, 32, 33, 30, 30, 32, 24, 25, 24, 22, 26, 26, 25, 26, 28, 23),
  Hips = c(32, 37, 36, 39, 33, 38, 42, 40, 37, 39, 35, 37, 37, 34, 38, 37, 38, 37, 40, 35)),
  .Names = c("Chest", "Waist", "Hips"), row.names = c(NA, -20L), class = "data.frame")
body
```

```
   Chest Waist Hips
1     34    30   32
2     37    32   37
3     38    30   36
4     36    33   39
5     38    29   33
6     43    32   38
7     40    33   42
8     38    30   40
9     40    30   37
10    41    32   39
11    36    24   35
12    36    25   37
13    34    24   37
14    33    22   34
15    36    26   38
16    37    26   37
17    34    25   38
18    36    26   37
19    38    28   40
20    35    23   35
```

```
colMeans(body); diag(round(var(body), digits = 2))
```

```
Chest Waist  Hips
37.00 28.00 37.05
```

```
Chest Waist  Hips
 6.63 12.53  5.94
```

```
round(var(body), digits = 2)
```

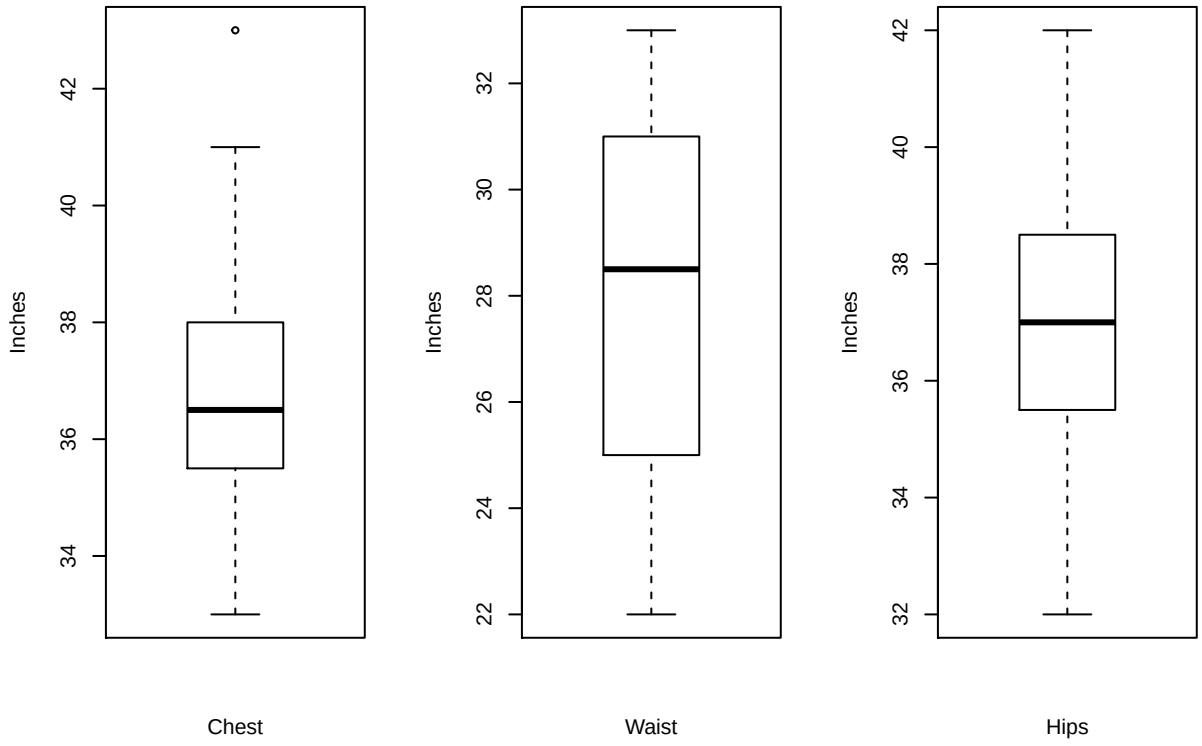
```
      Chest Waist Hips
Chest  6.63  6.37 3.00
Waist  6.37 12.53 3.58
Hips   3.00  3.58 5.94
```

```
round(cor(body), digits = 2)
```

```
      Chest Waist Hips
Chest  1.00  0.70 0.48
Waist  0.70  1.00 0.41
Hips   0.48  0.41 1.00
```

Figure 12.1

```
attach(body)
par(mfrow = c(1,3))
boxplot(Chest, ylab = "Inches", xlab = "Chest")
boxplot(Waist, ylab = "Inches", xlab = "Waist")
boxplot(Hips,  ylab = "Inches", xlab = "Hips")
```




```
detach(body)
```

Figure 12.2

```
# function for plotting histograms on main diagonal of pairs plot
panel.hist <- function(x, ...)
{
    usr <- par("usr"); on.exit(par(usr))
    par(usr = c(usr[1:2], 0, 1.5) )
    h <- hist(x, plot = FALSE)
    breaks <- h$breaks
    nB <- length(breaks)
    y <- h$counts
    y <- y/max(y)
    rect(breaks[-nB], 0, breaks[-1], y, col="cyan", ...)
}

pairs(body, diag.panel = panel.hist)
```

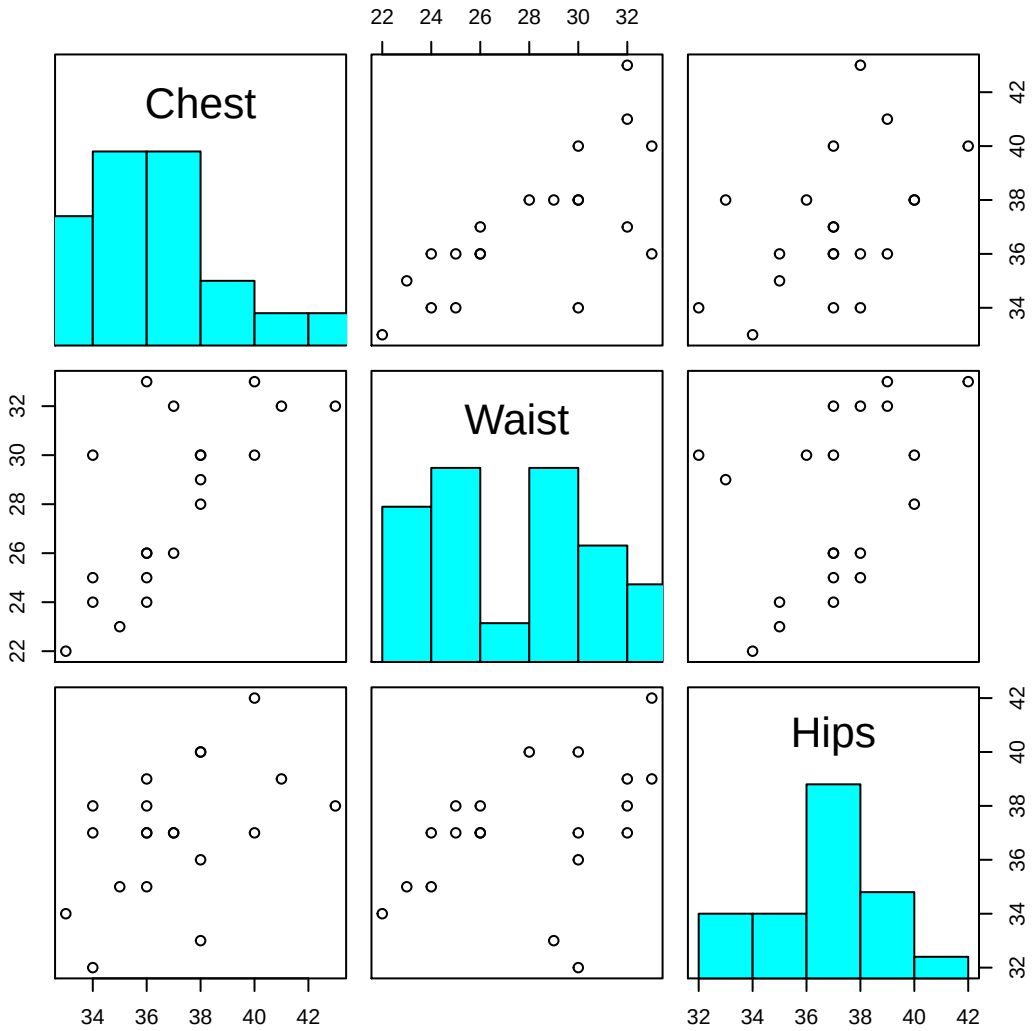


Figure 12.3

```
# bivariate normal density
x <- seq(-4, 4, length = 100)
y <- seq(-4, 4, length = 100)
mu1 <- 0
mu2 <- 0
sig1 <- 1
sig2 <- 1
rho <- 0.5

bvn <- function(x, y, mu1, mu2, sig1, sig2, rho) {
    denom <- 2*pi*sig1*sig2*sqrt(1-rho*rho)
    num <- ((x-mu1)/sig1)^2 - 2*rho*((x-mu1)/sig1)*((y-mu2)/sig2) + ((y-mu2)/sig2)^2
    num <- exp(-1/(2*(1-rho^2))*num)
    den <- num / denom
}
# mvden will be a 100x100 matrix:
mvden <- outer(x, y, FUN = bvn, mu1, mu2, sig1, sig2, rho)

library(lattice)
wireframe(mvden, xlab = expression(x[1]), ylab = expression(x[2]),
          zlab = list(expression(f(x[1], x[2])), rot = 90),
          screen = list(z = 25, x = -75, y = 0), zoom = 0.962,
          scales = list(distance = 0.5, col = 1),
          par.settings = list(axis.line = list(col = "transparent"))
)
```

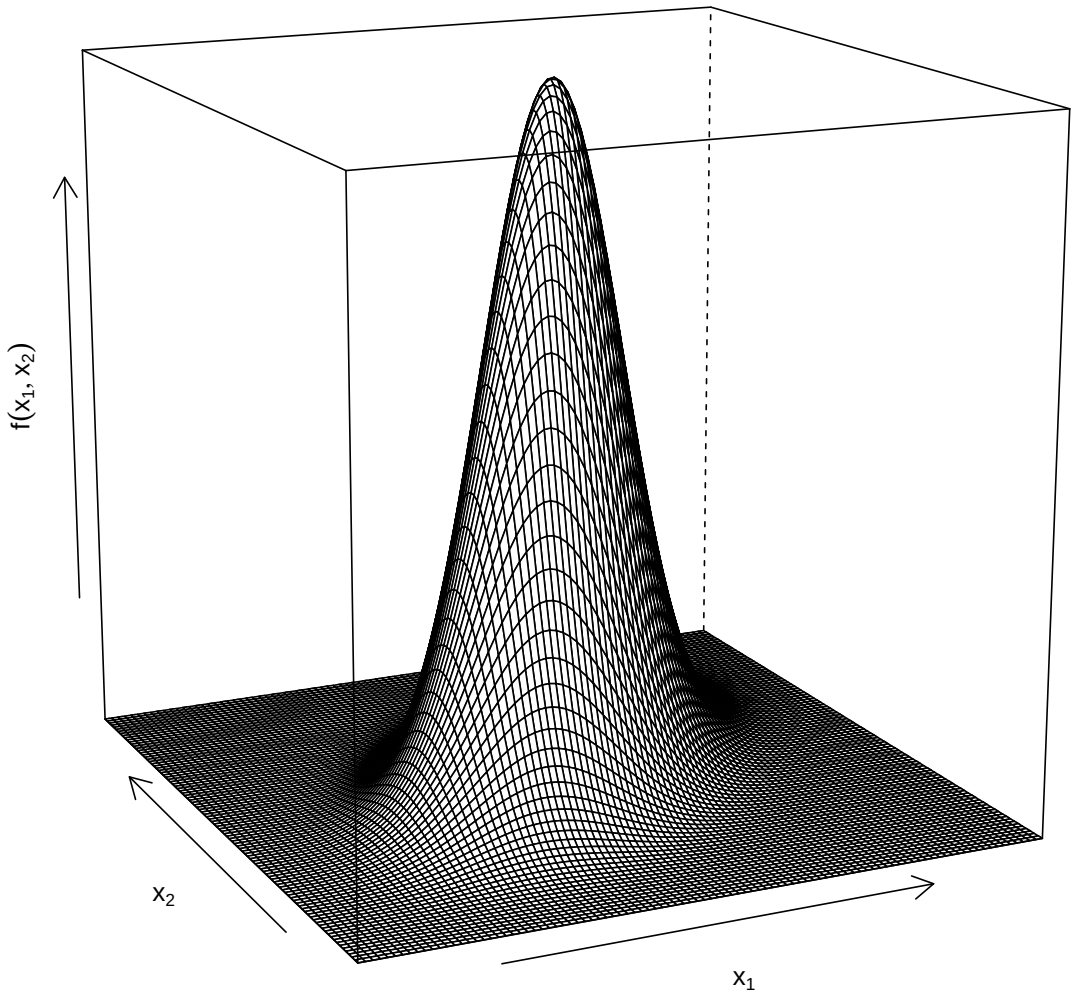
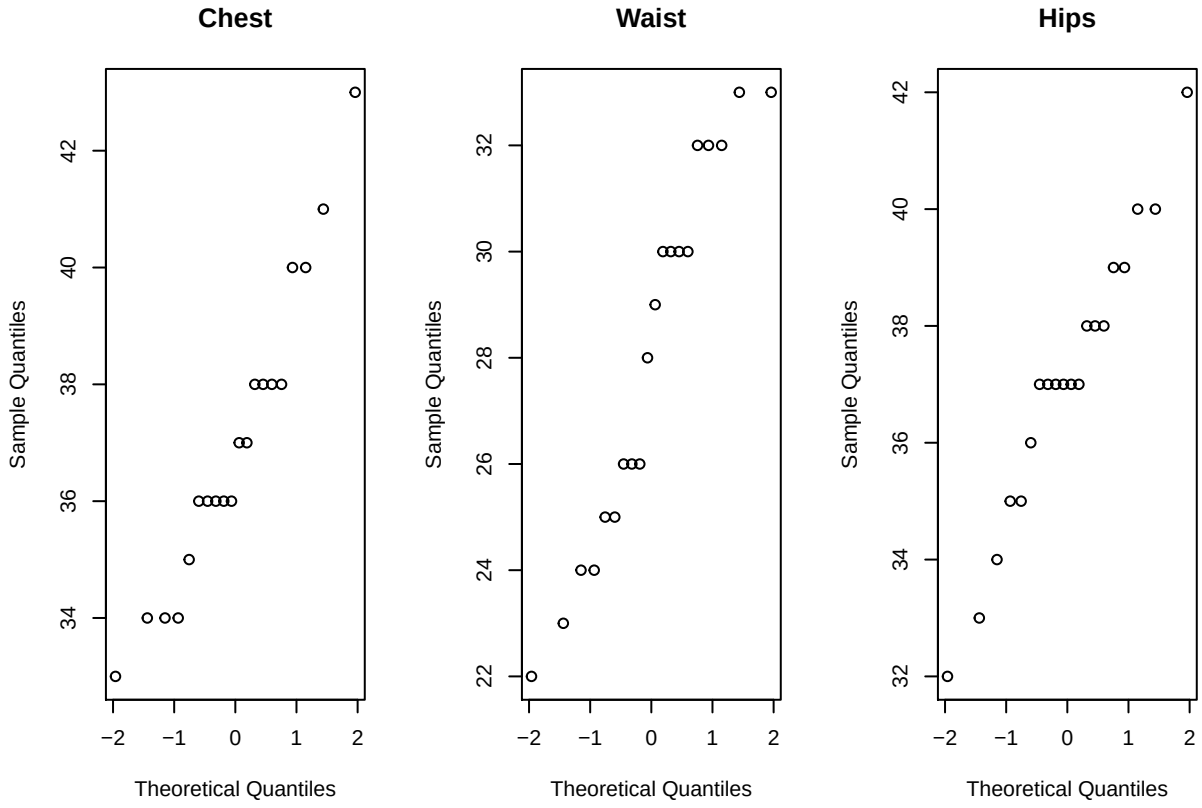

f(x1, x2)
x2
x1

Figure 12.4

```
attach(body)
par(mfrow=c(1,3))
qqnorm(Chest, main = "Chest")
qqnorm(Waist, main = "Waist")
qqnorm(Hips,  main = "Hips")
```




```
detach(body)
```

Figure 12.5

```
chisplot <- function(x) {
    if (!is.matrix(x)) stop("x is not a matrix")
    n <- nrow(x)
    p <- ncol(x)
    xbar <- apply(x, 2, mean)
    S <- var(x)
    S <- solve(S)
    index <- (1 : n) / (n + 1)
    xcent <- t(t(x) - xbar)
    di <- apply(xcent, 1, function(x,S) x %*% S %*% x, S)

    quant <- qchisq(index, p)
    plot(quant, sort(di),
         ylab = "Ordered distances",
         xlab = "Chi-square quantile",
         lwd = 2, pch = 1)
}
chisplot(as.matrix(body)); abline(0,1)
```

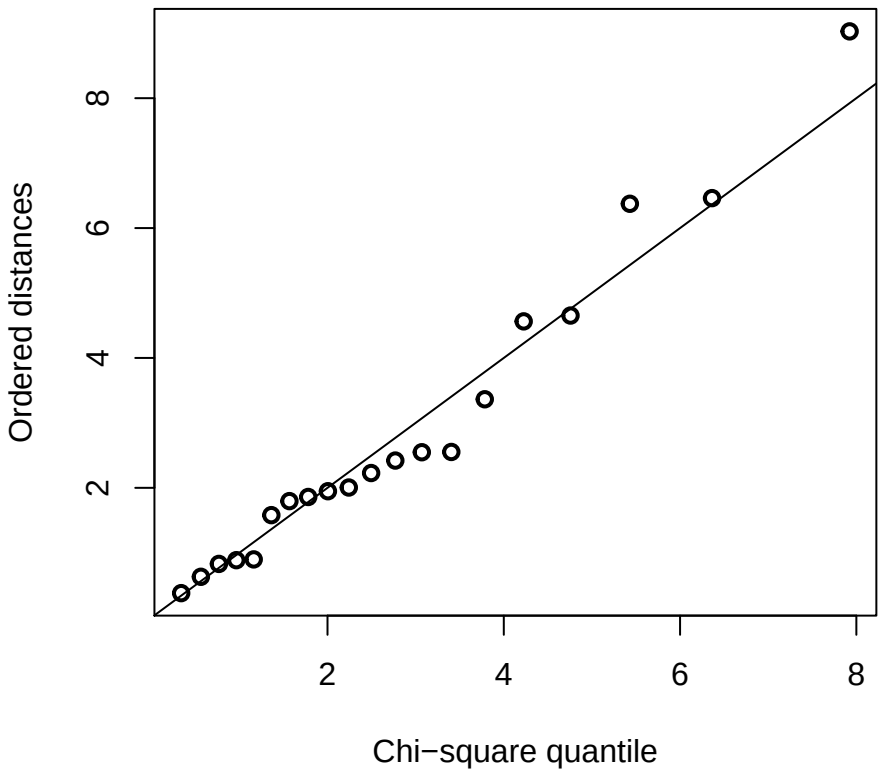


Table 12.2: Data on Paint Sprayers

```
paint <- read.table("data/paint.txt", header = TRUE, sep = '\t')

str(paint)
```

```
## 'data.frame':    102 obs. of  6 variables:
##  $ Haemo : num  13.4 14.6 13.5 15 14.6 14 16.4 14.8 15.2 15.5 ...
##  $ PCV   : int  39 46 42 46 44 44 49 44 46 48 ...
##  $ WBC   : int  4100 5000 4500 4600 5100 4900 4300 4400 4100 8400 ...
##  $ Lympho: int  14 15 19 23 17 20 21 16 27 34 ...
##  $ Neutro: int  25 30 21 16 31 24 17 26 13 42 ...
##  $ Lead  : int  17 20 18 18 19 19 18 29 27 36 ...
```

```
head(paint, n = 20)
```

```
##    Haemo PCV  WBC Lympho Neutro Lead
## 1   13.4  39 4100     14     25   17
## 2   14.6  46 5000     15     30   20
## 3   13.5  42 4500     19     21   18
## 4   15.0  46 4600     23     16   18
## 5   14.6  44 5100     17     31   19
## 6   14.0  44 4900     20     24   19
## 7   16.4  49 4300     21     17   18
## 8   14.8  44 4400     16     26   29
## 9   15.2  46 4100     27     13   27
## 10  15.5  48 8400     34     42   36
## 11  15.2  47 5600     26     27   22
## 12  16.9  50 5100     28     17   23
## 13  14.8  44 4700     24     20   23
## 14  16.2  45 5600     26     25   19
## 15  14.7  43 4000     23     13   17
## 16  14.7  42 3400      9     22   13
## 17  16.5  45 5400     18     32   17
## 18  15.4  45 6900     28     36   24
## 19  15.1  45 4600     17     29   17
## 20  14.2  46 4200     14     25   28
```

Figure 12.6

```
par(mfrow = c(2,3))
qqnorm(paint[, 1], main = "Haemo")
qqnorm(paint[, 2], main = "PCV")
qqnorm(paint[, 3], main = "WBC")
qqnorm(paint[, 4], main = "Lympho")
qqnorm(paint[, 5], main = "Neutro")
qqnorm(paint[, 6], main = "Lead")
```

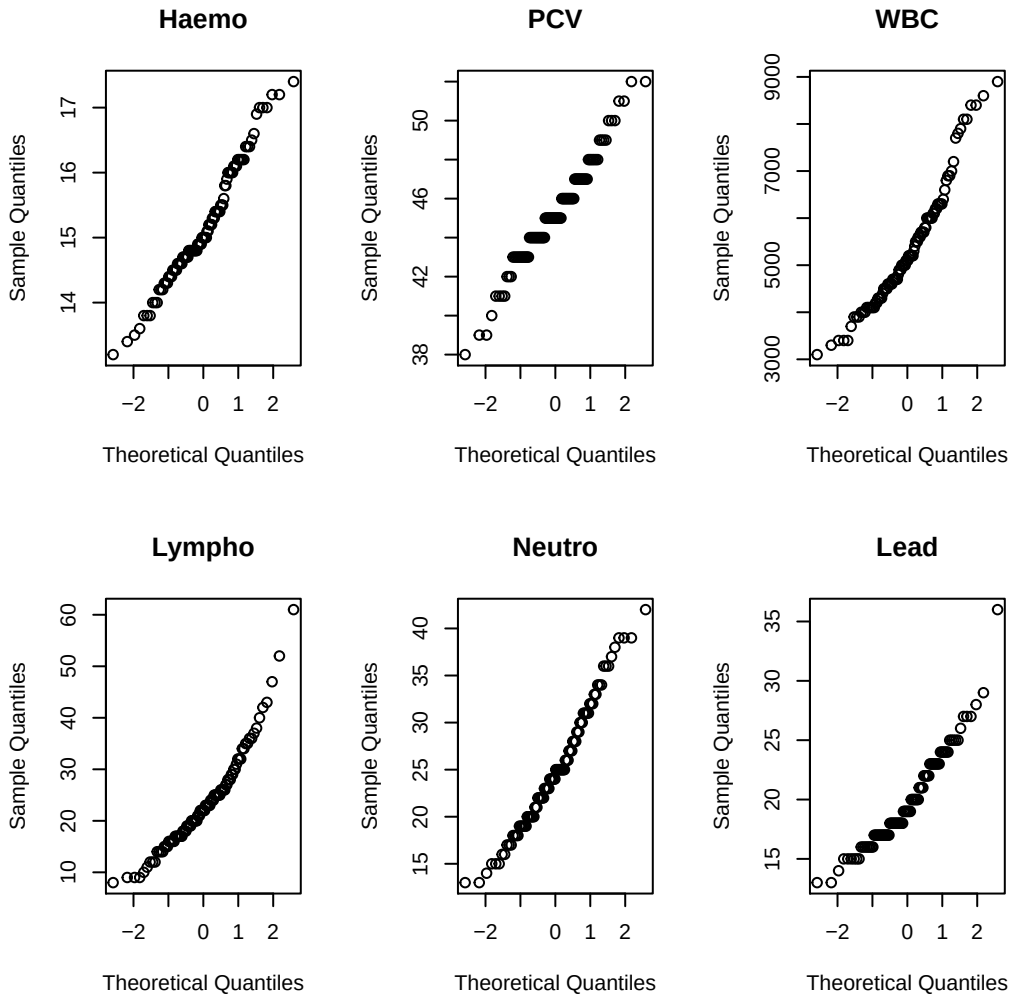


Figure 12.7

```
chisplot(as.matrix(paint))
abline(0,1)
```

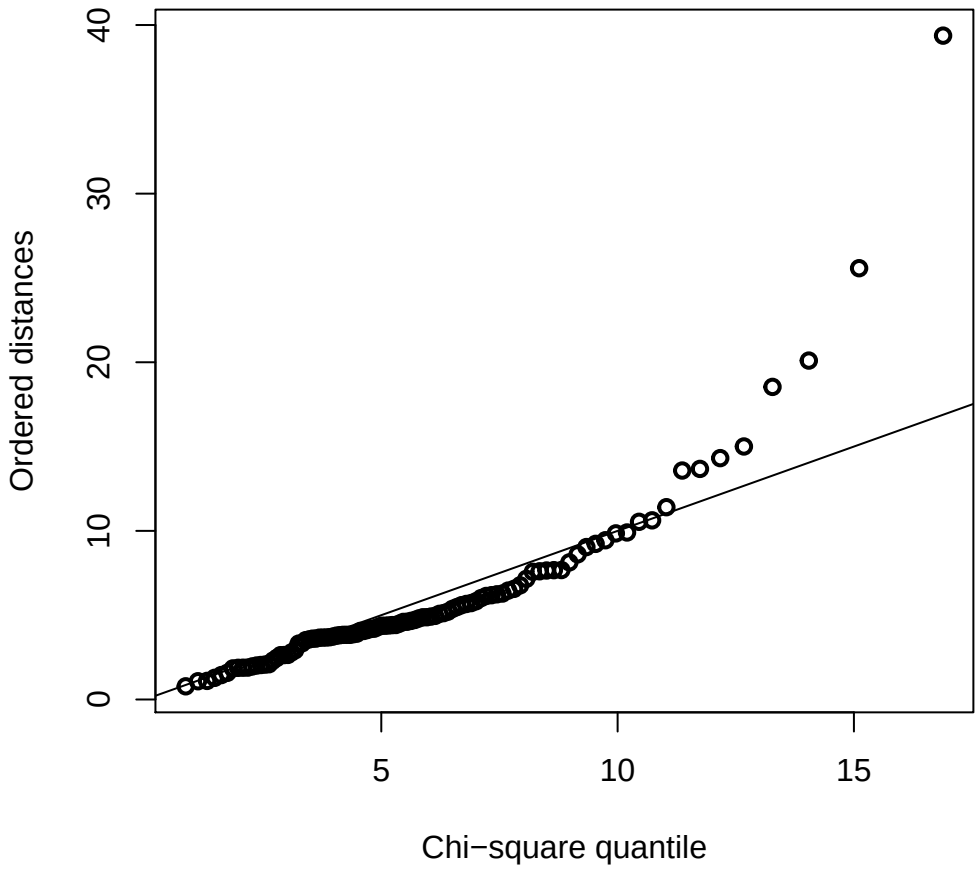


Figure 12.8

```
chisplot(as.matrix(log(paint)))
abline(0,1)
```