A bootstrap sample of the returns was drawn with the following R code. The returns are in the matrix dat and yboot is a bootstrap sample chosen by taking a random sample of the rows of dat, with replacement of course.

```
yboot = dat[sample((1:n), n, replace = TRUE), ]
```

□

7.12 Bibliographic Notes

The multivariate central limit theorem for the MLE is stated precisely and proved in textbooks on asymptotic theory such as Lehmann (1999) and van der Vaart (1998). The multivariate skewed t-distribution is in Azzalini and Capitanio (2003) and Azzalini (2014).

7.13 R Lab

7.13.1 Equity Returns

This section uses the data set berndtInvest on the book's web site and taken originally from R's fEcofin package. This data set contains monthly returns from January 1, 1987, to December 1, 1987, on 16 equities. There are 18 columns. The first column is the date and the last is the risk-free rate.

In the lab we will only use the first four equities. The following code computes the sample covariance and correlation matrices for these returns.

```
berndtInvest = read.csv("berndtInvest.csv")
Berndt = as.matrix(berndtInvest[, 2:5])
cov(Berndt)
cor(Berndt)
```

If you wish, you can also plot a scatterplot matrix with the following R code.

```
pairs(Berndt)
```

Problem 1 *Suppose the four variables being used are denoted by* $X_1, \ldots, X_4$*. Use the sample covariance matrix to estimate the variance of* $0.5X_1 + 0.3X_2 + 0.2X_3$*. (Useful* R *facts: "*`t(a)`*" is the transpose of a vector or matrix* `a` *and "*`a %*% b`*" is the matrix product of* `a` *and* `b`*.)*

Fit a multivariate-t model to the data using the function cov.trob in the MASS package. This function computes the MLE of the mean and covariance matrix with a fixed value of ν. To find the MLE of ν, the following code computes the profile log-likelihood for ν.

```
library(MASS)  # needed for cov.trob
library(mnormt)  # needed for dmt
df = seq(2.5, 8, 0.01)
n = length(df)
loglik_profile = rep(0, n)
for(i in 1:n)
{
  fit = cov.trob(Berndt, nu = df[i])
  mu = as.vector(fit$center)
  sigma = matrix(fit$cov, nrow = 4)
  loglik_profile[i] = sum(log(dmt(Berndt, mean = fit$center,
     S= f it$cov, df = df[i])))
}
```

Problem 2 *Using the results produced by the code above, find the MLE of ν and a 90 % profile likelihood confidence interval for ν. Include your* R *code with your work. Also, plot the profile log-likelihood and indicate the MLE and the confidence interval on the plot.*

Section 7.13.3 demonstrates how the MLE for a multivariate t-model can be fit directly with the `optim` function, rather than profile likelihood.

7.13.2 Simulating Multivariate *t*-Distributions

The following code generates and plots four bivariate samples. Each sample has univariate marginals that are standard t_3-distributions. However, the dependencies are different.

```
library(MASS)  # need for mvrnorm
par(mfrow=c(1,4))
N =  2500
nu = 3

set.seed(5640)
cov=matrix(c(1, 0.8, 0.8, 1), nrow = 2)
x= mvrnorm(N, mu = c(0, 0), Sigma = cov)
w = sqrt(nu / rchisq(N, df = nu))
x = x * cbind(w, w)
plot(x, main = "(a)")

set.seed(5640)
cov=matrix(c(1, 0.8, 0.8, 1),nrow = 2)
x= mvrnorm(N, mu = c(0, 0), Sigma = cov)
w1 = sqrt(nu / rchisq(N, df = nu))
w2 = sqrt(nu / rchisq(N, df = nu))
x = x * cbind(w1, w2)
plot(x, main = "(b)")
```

```
set.seed(5640)
cov=matrix(c(1, 0, 0, 1), nrow = 2)
x= mvrnorm(N, mu = c(0, 0), Sigma = cov)
w1 = sqrt(nu / rchisq(N, df = nu))
w2 = sqrt(nu / rchisq(N, df = nu))
x = x * cbind(w1, w2)
plot(x, main = "(c)")

set.seed(5640)
cov=matrix(c(1, 0, 0, 1), nrow = 2)
x= mvrnorm(N, mu = c(0, 0), Sigma = cov)
w = sqrt(nu / rchisq(N, df = nu))
x = x * cbind(w, w)
plot(x, main = "(d)")
```

Note the use of these R commands: `set.seed` to set the seed of the random number generator, `mvrnorm` to generate multivariate normally distributed vectors, `rchisq` to generate χ^2-distributed random numbers, `cbind` to bind together vectors as the columns of a matrix, and `matrix` to create a matrix from a vector. In R, "`a * b`" is elementwise multiplication of same-size matrices `a` and `b`, and "`a %*% b`" is matrix multiplication of conforming matrices `a` and `b`.

Problem 3 *Which sample has independent variates? Explain your answer.*

Problem 4 *Which sample has variates that are correlated but do not have tail dependence? Explain your answer.*

Problem 5 *Which sample has variates that are uncorrelated but with tail dependence? Explain your answer.*

Problem 6* *Suppose that (X, Y) are the returns on two assets and have a multivariate t-distribution with degrees of freedom, mean vector, and covariance matrix*

$$\nu = 5, \quad \mu = \begin{pmatrix} 0.001 \\ 0.002 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0.10 & 0.03 \\ 0.03 & 0.15 \end{pmatrix}.$$

Then $R = (X + Y)/2$ is the return on an equally weighted portfolio of the two assets.

(a) *What is the distribution of R?*
(b) *Write an* R *program to generate a random sample of size 10,000 from the distribution of* R. *Your program should also compute the 0.01 upper quantile of this sample and the sample average of all returns that exceed this quantile. This quantile and average will be useful later when we study risk analysis.*

7.13.3 Fitting a Bivariate t-Distribution

When you run the R code that follows this paragraph, you will compute the MLE for a bivariate t-distribution fit to CRSP returns data. A challenge when fitting a multivariate distribution is enforcing the constraint that the scale matrix (or the covariance matrix) must be positive definite. One way to meet this challenge is to let the scale matrix be $A^T A$, where A is an upper triangular matrix. (It is easy to show that $A^T A$ is positive semidefinite if A is any square matrix. Because a scale or covariance matrix is symmetric, only the entries on and above the main diagonal are free parameters. In order for A to have the same number of free parameters as the covariance matrix, we restrict A to be upper triangular.)

```
library(mnormt)
data(CRSPday, package = "Ecdat")
Y = CRSPday[ , c(5, 7)]
loglik = function(par)
{
mu = par[1:2]
A = matrix(c(par[3], par[4], 0, par[5]), nrow = 2, byrow = T)
scale_matrix = t(A) %*% A
df = par[6]
-sum(log(dmt(Y, mean = mu, S = scale_matrix, df = df)))
}
A = chol(cov(Y))
start = as.vector(c(apply(Y, 2, mean),
   A[1, 1], A[1, 2], A[2, 2], 4))
fit_mvt = optim(start, loglik, method = "L-BFGS-B",
   lower = c(-0.02, -0.02, -0.1, -0.1, -0.1, 2),
   upper = c(0.02, 0.02, 0.1, 0.1, 0.1, 15), hessian = T)
```

Problem 7* *Let* $\boldsymbol{\theta} = (\mu_1, \mu_2, A_{1,1}, A_{1,2}, A_{2,2}, \nu)$, *where* μ_j *is the mean of the* j*th variable,* $A_{1,1}$, $A_{1,2}$, *and* $A_{2,2}$ *are the nonzero elements of* A, *and* ν *is the degrees-of-freedom parameter.*

(a) What does the code `A = chol(cov(Y))` *do?*
(b) Find $\widehat{\boldsymbol{\theta}}_{\mathrm{ML}}$, *the MLE of* $\boldsymbol{\theta}$.
(c) Find the Fisher information matrix for $\boldsymbol{\theta}$. *(Hint: The Hessian is part of the object* `fit_mvt`. *Also, the* R *function* `solve` *will invert a matrix.)*
(d) Find the standard errors of the components of $\widehat{\boldsymbol{\theta}}_{\mathrm{ML}}$ *using the Fisher information matrix.*
(e) Find the MLE of the covariance matrix of the returns.
(f) Find the MLE of ρ, *the correlation between the two returns (*Y_1 *and* Y_2*).*

7.14 Exercises

1. Suppose that $E(X) = 1$, $E(Y) = 1.5$, $\text{Var}(X) = 2$, $\text{Var}(Y) = 2.7$, and $\text{Cov}(X,Y) = 0.8$.
   (a) What are $E(0.2X + 0.8Y)$ and $\text{Var}(0.2X + 0.8Y)$?
   (b) For what value of w is $\text{Var}\{wX + (1-w)Y\}$ minimized? Suppose that X is the return on one asset and Y is the return on a second asset. Why would it be useful to minimize $\text{Var}\{wX + (1-w)Y\}$?
2. Let X_1, X_2, Y_1, and Y_2 be random variables.
   (a) Show that $\text{Cov}(X_1 + X_2, Y_1 + Y_2) = \text{Cov}(X_1, Y_1) + \text{Cov}(X_1, Y_2) + \text{Cov}(X_2, Y_1) + \text{Cov}(X_2, Y_2)$.
   (b) Generalize part (a) to an arbitrary number of X_is and Y_is.
3. Verify formulas (A.24)–(A.27).
4. (a) Show that

$$E\{X - E(X)\} = 0$$

   for any random variable X.
   (b) Use the result in part (a) and Eq. (A.31) to show that if two random variables are independent then they are uncorrelated.
5. Show that if X is uniformly distributed on $[-a, a]$ for any $a > 0$ and if $Y = X^2$, then X and Y are uncorrelated but they are not independent.
6. Verify the following results that were stated in Sect. 7.3:

$$E(\boldsymbol{w}^\mathsf{T}\boldsymbol{X}) = \boldsymbol{w}^\mathsf{T}\{E(\boldsymbol{X})\}$$

   and

$$\begin{aligned}\text{Var}(\boldsymbol{w}^\mathsf{T}\boldsymbol{X}) &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_i\, w_j\, \text{Cov}(X_i, X_j) \\ &= \text{Var}(\boldsymbol{w}^\mathsf{T}\boldsymbol{X})\boldsymbol{w}^\mathsf{T}\text{COV}(\boldsymbol{X})\boldsymbol{w}.\end{aligned}$$

7. Suppose $\mathbf{Y} = (Y_1, Y_2, Y_3)$ has covariance matrix

$$COV(\mathbf{Y}) = \begin{pmatrix} 1.0 & 0.9 & a \\ 0.9 & 1.0 & 0.9 \\ a & 0.9 & 1.0 \end{pmatrix}$$

   for some unknown value a. Use Eq. (7.7) and the fact that the variance of a random variable is always ≥ 0 to show that a cannot equal 0.