Homework 4

DATA130021 Financial Econometrics

Deng Qisheng

2021/04/12

Problem 1

Suppose that $\mathbb{E}(X) = 1$, $\mathbb{E}(Y) = 2$, $\text{Var}(X) = 2$, $\text{Var}(Y) = 2.7$ and $\text{Cov}(X, Y) = 0.8$.

(a) What are $E(0.2X + 0.8Y)$ and $\text{Var}(0.2X + 0.8Y)$?

Solution:

Recall the properties of expectation and variance, we have

$$\begin{aligned}
\mathbb{E}(0.2X + 0.8Y) &= 0.2\ \mathbb{E}(X) + 0.8\ \mathbb{E}(Y) \\
&= 0.2 \times 1 + 0.8 \times 1.5 \\
&= 1.4,
\end{aligned}$$

$$\begin{aligned}
\text{Var}(0.2X + 0.8Y) &= \text{Var}(0.2X) + \text{Var}(0.8Y) + 2\text{Cov}(0.2X, 0.8Y) \\
&= 0.2^2\text{Var}(X) + 0.8^2\text{Var}(Y) + 2 \times 0.2 \times 0.8\text{Cov}(X, Y) \\
&= 0.04 \times 2 + 0.64 \times 2.7 + 0.32 \times 0.8 \\
&= 2.064.
\end{aligned}$$

(b) For what value of ω is $\text{Var}(\omega X + (1-\omega)Y)$ 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 $\text{Var}(\omega X + (1-\omega)Y)$?

Solution:

Consider the formula of variance, then

$$\begin{aligned}
\text{Var}(\omega X + (1-\omega)Y) &= \text{Var}(\omega X) + \text{Var}((1-\omega)Y) + 2\text{Cov}(\omega X, (1-\omega)Y) \\
&= \omega^2\text{Var}(X) + (1-\omega)^2\text{Var}(Y) + 2\omega(1-\omega)\text{Cov}(X, Y) \\
&= 2\omega^2 + 2.7(1-\omega)^2 + 1.6\omega(1-\omega) \\
&= 3.1\omega^2 - 3.8\omega + 2.7 \\
&\approx 3.1(\omega - 0.6129)^2 + 1.5355.
\end{aligned}$$

Hence, $\text{Var}(\omega X + (1-\omega)Y)$ is minimized when $\omega \approx 0.6129$. If X is the return on one asset and Y is the return on a second asset, then value $\omega X + (1-\omega)Y$ can be seen as the return of a combination of two assets and value ω is the ratio of two assets. A optimized ω means that return of the portfolio has the minimized risk, which is measured by the variance $\text{Var}(\omega X + (1-\omega)Y)$.

Problem 2

Kendall's tau rank correlation between X and Y is 0.55. Both X and Y are positive. What is Kendall's tau between X and $1/Y$? What is Kendall's tau between $1/X$ and $1/Y$?

Solution:

From the assumption, for bivariate samples (X_1, Y_1) and (X_2, Y_2), we have

$$\rho_\tau(X, Y) = P\Big((X_1 - X_2)(Y_1 - Y_2) > 0\Big) - P\Big((X_1 - X_2)(Y_1 - Y_2) < 0\Big) = 0.55.$$

Hence, Kendall's tau between X and $1/Y$ is

$$\begin{aligned}
\rho_\tau(X, 1/Y) &= P\Big((X_1 - X_2)(1/Y_1 - 1/Y_2) > 0\Big) - P\Big((X_1 - X_2)(1/Y_1 - 1/Y_2) < 0\Big) \\
&= P\left(\frac{(X_1 - X_2)(Y_2 - Y_1)}{Y_1 Y_2} > 0\right) - P\left(\frac{(X_1 - X_2)(Y_2 - Y_1)}{Y_1 Y_2} < 0\right) \\
&= P\Big((X_1 - X_2)(Y_2 - Y_1) > 0\Big) - P\Big((X_1 - X_2)(Y_2 - Y_1) < 0\Big) \\
&= P\Big((X_1 - X_2)(Y_1 - Y_2) < 0\Big) - P\Big((X_1 - X_2)(Y_1 - Y_2) > 0\Big) \\
&= -\rho_\tau(X, Y) \\
&= -0.55.
\end{aligned}$$

Kendall's tau between $1/X$ and $1/Y$ is

$$\begin{aligned}
\rho_\tau(1/X, 1/Y) &= P\Big((1/X_1 - 1/X_2)(1/Y_1 - 1/Y_2) > 0\Big) - P\Big((1/X_1 - 1/X_2)(1/Y_1 - 1/Y_2) < 0\Big) \\
&= P\left(\frac{(X_2 - X_1)(Y_2 - Y_1)}{X_1 X_2 Y_1 Y_2} > 0\right) - P\left(\frac{(X_2 - X_1)(Y_2 - Y_1)}{X_1 X_2 Y_1 Y_2} < 0\right) \\
&= P\Big((X_2 - X_1)(Y_2 - Y_1) > 0\Big) - P\Big((X_2 - X_1)(Y_2 - Y_1) < 0\Big) \\
&= P\Big((X_1 - X_2)(Y_1 - Y_2) > 0\Big) - P\Big((X_1 - X_2)(Y_1 - Y_2) < 0\Big) \\
&= \rho_\tau(X, Y) \\
&= 0.55.
\end{aligned}$$

Problem 3

Show that an Archimedean copula with generator function $\phi(u) = -\log(u)$ is equal to the independence copula C_0. Does the same hold when the natural logarithm is replaced by the common logarithm, i.e., $\phi(u) = -\log_{10}(u)$?

Proof:

For natural logarithm, an Archimedean copula with generator function $\phi(u) = -\log(u)$ can be written as

$$\begin{aligned}
C_{\text{natural}}(u_1, u_2) &= \phi^{-1}\Big(\phi(u_1) + \phi(u_2)\Big) \\
&= e^{-\big(-\log(u_1) - \log(u_2)\big)} \\
&= e^{\log(u_1) + \log(u_2)} \\
&= u_1 u_2 \\
&= C_0,
\end{aligned}$$

which is equal to the independence copula C_0. For common logarithm, an Archimedean copula with generator function $\phi(u) = -\log_a(u)$ can be written as

$$\begin{aligned} C_{\text{common}}(u_1, u_2) &= \phi^{-1}\Big(\phi(u_1) + \phi(u_2)\Big) \\ &= a^{-\big(-\log_a(u_1) - \log_a(u_2)\big)} \\ &= a^{\log_a(u_1) + \log_a(u_2)} \\ &= u_1 u_2 \\ &= C_0, \end{aligned}$$

which is also equal to the independence copula C_0. ■

Problem 4

Suppose $\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}_2\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}\right)$. Find the conditional distribution of $X|Y = y$.

Solution:

Recall the density function of bivariate normal distribution

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2}) + (\frac{y-\mu_2}{\sigma_2})^2\right]}, \quad x, y \in (-\infty, +\infty).$$

Then we can find the marginal density function of Y as

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f_{X,Y}(x, y)\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2}) + (\frac{y-\mu_2}{\sigma_2})^2\right]}\mathrm{d}x \\ &= \frac{1}{2\pi\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2}) + \rho^2(\frac{y-\mu_2}{\sigma_2})^2\right] - \frac{1}{2}(\frac{y-\mu_2}{\sigma_2})^2} \mathrm{d}\Big(\frac{x-\mu_1}{\sigma_1}\Big) \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}(\frac{y-\mu_2}{\sigma_2})^2} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{u}{\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}}(\frac{y-\mu_2}{\sigma_2})\right]^2} \mathrm{d}\Big(\frac{u}{\sqrt{1-\rho^2}}\Big) \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}(\frac{y-\mu_2}{\sigma_2})^2} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2}\mathrm{d}v \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}(\frac{y-\mu_2}{\sigma_2})^2} \times 1 \\ &= \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}(\frac{y-\mu_2}{\sigma_2})^2}, \quad y \in (-\infty, +\infty). \end{aligned}$$

Similarly, we know that

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}(\frac{x-\mu_1}{\sigma_1})^2}, \quad x \in (-\infty, +\infty).$$

which means that the marginal distributions of bivariate normal distribution are univariate normal distributions, that is,

$$\begin{aligned} X &\sim \mathcal{N}(\mu_1, \sigma_1^2), \\ Y &\sim \mathcal{N}(\mu_2, \sigma_2^2). \end{aligned}$$

Hence, we can find the conditional density function of $X|Y = y$

$$
\begin{aligned}
f_{X|Y}\big(X|Y = y\big) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\
&= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2}) + (\frac{y-\mu_2}{\sigma_2})^2\right]} \times \sqrt{2\pi}\sigma_2 e^{\frac{1}{2}(\frac{y-\mu_2}{\sigma_2})^2} \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_1}{\sigma_1})^2 - 2\rho(\frac{x-\mu_1}{\sigma_1})(\frac{y-\mu_2}{\sigma_2}) + \rho^2(\frac{y-\mu_2}{\sigma_2})^2\right]} \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}} e^{-\frac{1}{2\sigma_1^2(1-\rho^2)}\left[(x-\mu_1) - \frac{\rho\sigma_1}{\sigma_2}(y-\mu_2)\right]^2} \\
&= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad x \in \big(-\infty, +\infty\big).
\end{aligned}
$$

where $\mu = \mu_1 + \frac{\rho\sigma_1}{\sigma_2}(y - \mu_2)$ and $\sigma = \sigma_1\sqrt{1-\rho^2}$. Hence, the conditional distribution of $X|Y = y$ is

$$
X|Y = y \ \sim \ \mathcal{N}\bigg(\mu_1 + \frac{\rho\sigma_1}{\sigma_2}(y - \mu_2), \sigma_1^2\big(1-\rho^2\big)\bigg).
$$

Problem 5

Let U_1, U_2 be i.i.d. $U[0,1]$ and define

$$
X_1 = \sqrt{-2\log U_1}\cos\big(2\pi U_1\big) \quad \text{and} \quad X_2 = \sqrt{-2\log U_2}\cos\big(2\pi U_2\big).
$$

Then, X_1 and X_2 are independent $\mathcal{N}\big(0,1\big)$ random variables. Use this method and `runif` function in R to:

(a) Write a function `my_rnorm1` to generate i.i.d. $\mathcal{N}\big(0,1\big)$ with sample size n.

Solution:

Codes of function `my_rnorm1` are shown below, where n is sample size.

```
my_rnorm1 <- function (n) {
    U <- runif(2 * n)
    X <- sqrt(-2 * log(U[1 : n])) * cos(2 * pi * U[(n + 1) : (2 * n)])
    return (X)
}
```

(b) Write a function `my_rnorm2` to generate i.i.d. $\mathcal{N}\big(\mu,\sigma^2\big)$ with sample size n.

Solution:

Codes of function `my_rnorm2` are shown below, where n is sample size, μ is mean and σ^2 is variance.

```
my_rnorm2 <- function (n, mu, sigma) {
    U <- runif(2 * n)
    X <- mu + sigma * sqrt(-2 * log(U[1 : n])) * cos(2 * pi * U[(n + 1) : (2 * n)])
    return (X)
}
```

(c) Generate some samples in (a) and (b) to test normality and paste all of your codes in a tidy way.

Solution:

Generate 1000 samples using `my_rnorm1` and `my_rnorm2`, respectively. Use KS, CVM and JB methods to test the normality. We can conclude that samples generated by methods above satisfy the normal distribution.

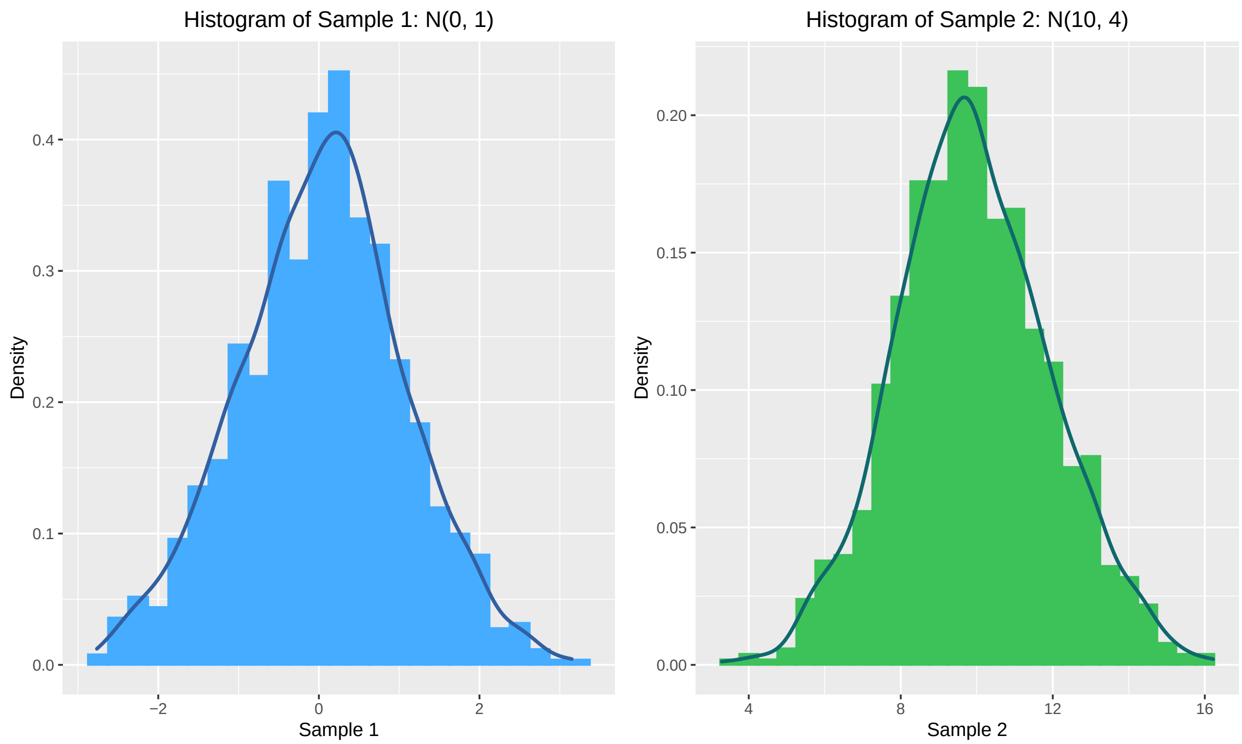
```r
# simulation
set.seed(2021)
sample.1 <- my_rnorm1(1000)
sample.2 <- my_rnorm2(1000, 10, 2)

# histogram
library(ggplot2)
library(gridExtra)

# sample 1
sample.1.hist <- ggplot(data.frame(sample.1), aes(x=sample.1, y=..density..)) +
    geom_histogram(binwidth=0.25, fill="#45ACFF", color="#45ACFF") +
    geom_density(color="#335FA0", size=1) +
    labs(title="Histogram of Sample 1: N(0, 1)", x="Sample 1", y="Density") +
    theme(plot.title=element_text(hjust=0.5))

# sample 2
sample.2.hist <- ggplot(data.frame(sample.2), aes(x=sample.2, y=..density..)) +
    geom_histogram(binwidth=0.5, fill="#3DC159", color="#3DC159") +
    geom_density(color="#10686C", size=1) +
    labs(title="Histogram of Sample 2: N(10, 4)", x="Sample 2", y="Density") +
    theme(plot.title=element_text(hjust=0.5))

# plot
grid.arrange(sample.1.hist, sample.2.hist, layout_matrix=rbind(c(1, 2)))
```




```r
# normality test
library(goftest)
library(tseries)
```

```r
# sample 1
ks.result.1 <- ks.test(sample.1, "pnorm", mean=mean(sample.1), sd=sd(sample.1))
cvm.result.1 <- cvm.test(sample.1, "pnorm", mean=mean(sample.1), sd=sd(sample.1))
jarque.bera.result.1 <- jarque.bera.test(sample.1)

# sample 2
ks.result.2 <- ks.test(sample.2, "pnorm", mean=mean(sample.2), sd=sd(sample.2))
cvm.result.2 <- cvm.test(sample.2, "pnorm", mean=mean(sample.2), sd=sd(sample.2))
jarque.bera.result.2 <- jarque.bera.test(sample.2)

# result
knitr::kable(
    x=data.frame(
        c("Kolmogorov-Smirnov Test", "Cramer-von Mises Test", "Jarque-Bera Test"),
        c(format(ks.result.1$p.value, scientific=F, digits=3, nsmall=3),
          format(cvm.result.1$p.value, scientific=F, digits=3, nsmall=3),
          format(jarque.bera.result.2$p.value, scientific=F, digits=3, nsmall=3)),
        c(format(ks.result.2$p.value, scientific=F, digits=3, nsmall=3),
          format(cvm.result.2$p.value, scientific=F, digits=3, nsmall=3),
          format(jarque.bera.result.2$p.value, scientific=F, digits=3, nsmall=3))),
    caption="Normality Test",
    col.names=c("P-value", "Sample 1: N(0, 1)", "Sample 2: N(10, 4)"),
    align=c("c", "c", "c"))
```

Table 1: Normality Test

P-value	Sample 1: N(0, 1)	Sample 2: N(10, 4)
Kolmogorov-Smirnov Test	0.709	0.304
Cramer-von Mises Test	0.721	0.506
Jarque-Bera Test	0.356	0.356

Problem 6

(X_i, Y_i) are i.i.d random vectors where all X_i satisfy $\mathcal{N}(0,1)$, all Y_i satisfy t_6 and the copula of (X_i, Y_i) is the t-copula with degree of freedom 4 and correlation 0.5. Simualte the random vectors and plot them in a scattor plot.

Solution:

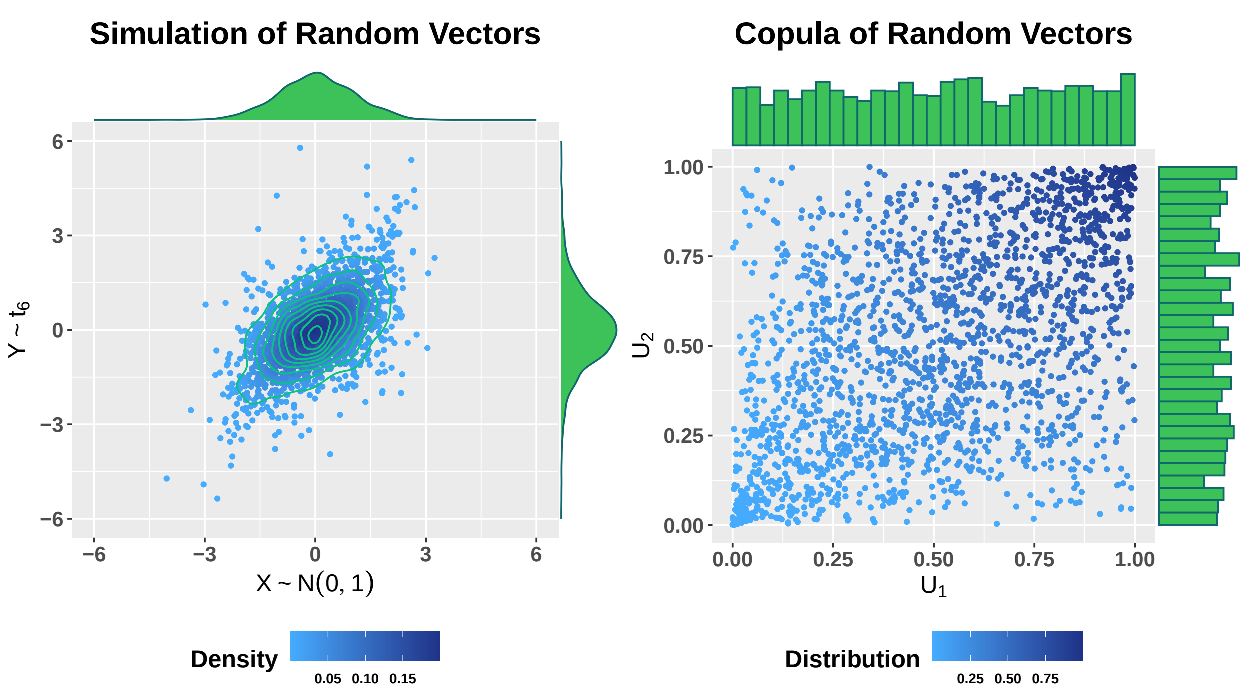
```r
# packages
library(copula)
library(ggplot2)
library(ggExtra)
library(gridExtra)

# simulation
set.seed(2021)
my_mvdc <- mvdc(copula=tCopula(param=0.5, dim=2, df=4), margins=c("norm", "t"),
    paramMargins = list(list(mean=0, sd=1), list(df=6)))
samples <- rMvdc(n=2000, mvdc=my_mvdc)
samples.cop <- cbind(pnorm(samples[, 1], mean=0, sd=1), pt(samples[, 2], df=6))
samples.density <- dMvdc(samples, my_mvdc)
samples.dist <- pMvdc(samples, my_mvdc)
```

```r
# random vectors
rv <- ggMarginal(
    ggplot(data.frame(samples), aes(x=X1, y=X2, color=samples.density)) +
        geom_point(size=1) + stat_density2d(color="#0BC286") + xlim(-6, 6) + ylim(-6, 6) +
        labs(title="Simulation of Random Vectors", color="Density",
             x=expression(X%~%N(0,1)), y=expression(Y%~%t[6])) +
        theme(plot.title=element_text(face="bold", size=18, hjust=0.5),
              axis.title=element_text(face="bold", size=14),
              axis.text=element_text(face="bold", size=12),
              legend.title=element_text(face="bold", size=14),
              legend.text=element_text(face="bold", size=8), legend.position="bottom") +
        scale_colour_gradient(high="#1E3388", low="#45ACFF"),
    type="density", fill="#3DC159", color="#10686C", size=8)

# copula
cop <- ggMarginal(
    ggplot(data.frame(samples.cop), aes(x=X1, y=X2, color=samples.dist)) +
        geom_point(size=1) +
        labs(title="Copula of Random Vectors", color="Distribution",
            x=expression(U[1]), y=expression(U[2])) +
        theme(plot.title=element_text(face="bold", size=18, hjust=0.5),
              axis.title=element_text(face="bold", size=14),
              axis.text=element_text(face="bold", size=12),
              legend.title=element_text(face="bold", size=14),
              legend.text=element_text(face="bold", size=8), legend.position="bottom") +
        scale_colour_gradient(high="#1E3388", low="#45ACFF"),
    type="histogram", fill="#3DC159", color="#10686C")

# plot
grid.arrange(rv, cop, layout_matrix=rbind(c(1, 2)))
```



Problem 7*

Use the JP Morgan and S&P 500 index daily return data from 2012/01/01 to 2018/12/31 to fit a normal copula and a t-copula by MLE. Make conclusion for your results.

Solutions:

This part is unfinished.

```
# packages
library(lubridate)

# data
jpm <- read.csv("./JPM.csv")
spx <- read.csv("./SPX.csv")
jpm$Date <- as.Date(jpm$Date)
spx$Date <- as.Date(spx$Date)
stock <- merge(jpm, spx, by="Date", all=F)
n <- dim(stock)[1]
Year <- as.factor(year(stock$Date))
Date <- stock$Date
returns <- cbind(Year, Date, rbind(c(0, 0), stock[2:n, -1] / stock[1:(n - 1), -1] - 1))

# plot
ggMarginal(
    ggplot(data.frame(returns), aes(x=JPM, y=SPX, colour=Year)) +
        geom_point(size=2) +
        labs(title="Returns of JPM and SPX", x="JPM", y="SPX", colour="Year") +
        theme(plot.title=element_text(face="bold", size=18, hjust=0.5),
              axis.title=element_text(face="bold", size=14),
              axis.text=element_text(face="bold", size=12),
              legend.title=element_text(face="bold", size=14),
              legend.text=element_text(face="bold", size=8), legend.position="bottom") +
        scale_colour_manual(values=c(
            "#A4D8FB", "#84C9FB", "#66B8FB", "#4EA4FB", "#4088DD", "#3571BE", "#335FA0")),
    type="histogram", fill="#3DC159", color="#10686C")
```