Homework 3

DATA130021 Financial Econometrics

Deng Qisheng

2021/03/19

Problem 1

(a) What is the kurtosis of a normal mixture distribution that is 95% $\mathcal{N}(0,1)$ and 5% $\mathcal{N}(0,10)$?

Solution:

Suppose $X_1 \sim \mathcal{N}(0,1)$, $X_2 \sim \mathcal{N}(0,10)$ and $Y \sim 0.95\mathcal{N}(0,1) + 0.05\mathcal{N}(0,10)$, then

$$\begin{aligned}
\mu = \mathbb{E}(Y) &= \mathbb{E}_X\Big[\mathbb{E}(Y|X)\Big] \\
&= P(X_1)\mathbb{E}(Y|X_1) + P(X_2)\mathbb{E}(Y|X_2) \\
&= 0.95 \times 0 + 0.05 \times 0 = 0,
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(Y^2) &= \mathbb{E}_X\Big[\mathbb{E}(Y^2|X)\Big] \\
&= P(X_1)\mathbb{E}(Y^2|X_1) + P(X_2)\mathbb{E}(Y^2|X_2) \\
&= P(X_1)\Big(\mathrm{Var}(Y|X_1) + \mathbb{E}^2(Y|X_1)\Big) + P(X_2)\Big(\mathrm{Var}(Y|X_2) + \mathbb{E}^2(Y|X_2)\Big) \\
&= 0.95 \times 1 + 0.05 \times 10 = 1.45,
\end{aligned}$$

$$\sigma^2 = \mathrm{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}^2(Y) = 1.45.$$

Hence, the kurtosis of Y is

$$\begin{aligned}
\kappa(Y) &= \mathbb{E}\left(\frac{Y-\mu}{\sigma}\right)^4 - 3 = \frac{1}{(\sigma^2)^2}\mathbb{E}(Y^4) - 3 = \frac{1}{(\sigma^2)^2}\mathbb{E}_X\Big[\mathbb{E}(Y^4|X)\Big] - 3 \\
&= \frac{1}{(\sigma^2)^2}\Big\{P(X_1)\mathbb{E}(Y^4|X_1) + P(X_2)\mathbb{E}(Y^4|X_2)\Big\} - 3 \\
&= \frac{1}{(\sigma^2)^2}\Big\{P(X_1)\Big(\mathrm{Var}(Y^2|X_1) + \mathbb{E}^2(Y^2|X_1)\Big) + P(X_2)\Big(\mathrm{Var}(Y^2|X_2) + \mathbb{E}^2(Y^2|X_2)\Big)\Big\} - 3 \\
&= \frac{1}{(\sigma^2)^2}\Big\{P(X_1)\Big(\mathrm{Var}(\chi_1^2) + \mathbb{E}^2(Y^2|X_1)\Big) + P(X_2)\Big(100\mathrm{Var}(\chi_1^2) + \mathbb{E}^2(Y^2|X_2)\Big)\Big\} - 3 \\
&= \frac{1}{(1.45)^2}\Big\{0.95 \times (2 + 1^2) + 0.05 \times (200 + 10^2)\Big\} - 3 \\
&\approx 5.49
\end{aligned}$$

(b) Find a formula for the kurtosis of a normal mixture distribution that is $100p\%$ $\mathcal{N}(0,1)$ and $100(1-p)\%$ $\mathcal{N}(0,\sigma^2)$, where p and σ are parameters. Your formula should give the kurtosis as a function of p and σ.

Solution:

Suppose $X_1 \sim \mathcal{N}(0,1)$, $X_2 \sim \mathcal{N}(0,\sigma^2)$ and $Y \sim p\mathcal{N}(0,1) + (1-p)\mathcal{N}(0,\sigma^2)$, then

$$\begin{aligned}\mathbb{E}(Y) &= \mathbb{E}_X\Big[\mathbb{E}(Y|X)\Big] \\ &= P(X_1)\mathbb{E}(Y|X_1) + P(X_2)\mathbb{E}(Y|X_2) \\ &= p \times 0 + (1-p) \times 0 = 0,\end{aligned}$$

$$\begin{aligned}\mathbb{E}(Y^2) &= \mathbb{E}_X\Big[\mathbb{E}(Y^2|X)\Big] \\ &= P(X_1)\mathbb{E}(Y^2|X_1) + P(X_2)\mathbb{E}(Y^2|X_2) \\ &= P(X_1)\Big(\mathrm{Var}(Y|X_1) + \mathbb{E}^2(Y|X_1)\Big) + P(X_2)\Big(\mathrm{Var}(Y|X_2) + \mathbb{E}^2(Y|X_2)\Big) \\ &= p \times 1 + (1-p) \times \sigma^2 = p + (1-p)\sigma^2,\end{aligned}$$

$$\mathrm{Var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}^2(Y) = p + (1-p)\sigma^2.$$

Hence, the kurtosis of Y is

$$\begin{aligned}\kappa(Y|p,\sigma^2) &= \mathbb{E}\left(\frac{Y-\mathbb{E}(Y)}{\sqrt{\mathrm{Var}(Y)}}\right)^4 - 3 = \frac{1}{\mathrm{Var}^2(Y)}\mathbb{E}(Y^4) - 3 = \frac{1}{\mathrm{Var}^2(Y)}\mathbb{E}_X\Big[\mathbb{E}(Y^4|X)\Big] - 3 \\ &= \frac{1}{\mathrm{Var}^2(Y)}\Big\{P(X_1)\mathbb{E}(Y^4|X_1) + P(X_2)\mathbb{E}(Y^4|X_2)\Big\} - 3 \\ &= \frac{1}{\mathrm{Var}^2(Y)}\Big\{P(X_1)\Big(\mathrm{Var}(Y^2|X_1) + \mathbb{E}^2(Y^2|X_1)\Big) + P(X_2)\Big(\mathrm{Var}(Y^2|X_2) + \mathbb{E}^2(Y^2|X_2)\Big)\Big\} - 3 \\ &= \frac{1}{\mathrm{Var}^2(Y)}\Big\{P(X_1)\Big(\mathrm{Var}(\chi_1^2) + \mathbb{E}^2(Y^2|X_1)\Big) + P(X_2)\Big(\sigma^4\mathrm{Var}(\chi_1^2) + \mathbb{E}^2(Y^2|X_2)\Big)\Big\} - 3 \\ &= \frac{p(2+1^2) + (1-p) \times (2\sigma^4 + \sigma^4)}{\Big(p + (1-p)\sigma^2\Big)^2} - 3 \\ &= \frac{3\Big(p + (1-p)\sigma^4\Big)}{\Big(p + (1-p)\sigma^2\Big)^2} - 3 \\ &= \frac{3p(1-p)(1-\sigma^2)^2}{\Big(p + (1-p)\sigma^2\Big)^2}.\end{aligned}$$

(c) Show that the kurtosis of the normal mixtures in part (b) can be made arbitrarily large by choosing p and σ appropriately. Find values of p and σ so that the kurtosis is 10000 or larger.

Proof:

From the assumption, we are going to solve the following inequality

$$f(p,\sigma^2) = \kappa(Y|p,\sigma^2) = \frac{3p(1-p)(1-\sigma^2)^2}{\Big(p + (1-p)\sigma^2\Big)^2} \geq 10000.$$

Consider the derivatives of p

$$\frac{\partial f\left(p,\sigma^2\right)}{\partial p} = \frac{3\left(1-\sigma^2\right)^2\left[\sigma^2 - p\left(1+\sigma^2\right)\right]}{\left(p+\left(1-p\right)\sigma^2\right)^3},$$

If we see σ^2 as a constant, then let the derivative be zero and we get

$$p_* = \frac{\sigma^2}{1+\sigma^2}.$$

Consider the second derivative at p_*

$$\left.\frac{\partial^2 f\left(p,\sigma^2\right)}{\partial p^2}\right|_{p=p_*} = -\frac{3\left(1-\sigma^2\right)^2\left(1+\sigma^2\right)^4}{8\sigma^6} < 0,$$

which means p_* is a maxima. When $p = p_*$, the kurtosis of Y is

$$\kappa\left(Y|p_*,\sigma^2\right) = \frac{3\left(1-\sigma^2\right)^2}{4\sigma^2} \geq M$$

Hence, we have

$$3\sigma^4 - \left(4M+6\right)\sigma^2 + 3 \geq 0.$$

Solve the inequality above and we get

$$\sigma^2 \geq \frac{2\sqrt{M^2+3M}+2M+3}{3} > 1 \quad \text{or} \quad \sigma^2 \leq \frac{-2\sqrt{M^2+3M}+2M+3}{3} < 1.$$

Let $M = 10000$ and we have

$$\sigma \geq 115.4787,\ p \approx 0.999925$$

or

$$0 < \sigma \leq 8.659605 \times 10^{-3},\ p \approx 7.498313 \times 10^{-5},$$

which makes $\kappa \geq 10000$. Hence, the kurtosis of the normal mixtures $0.95\mathcal{N}\left(0,1\right) + 0.05\mathcal{N}\left(0,10\right)$ can be made arbitrarily large by choosing p and σ appropriately. For example, we can choose $p = 0.999925$ and $\sigma = 116$. ■

(d) Let $M > 0$ be arbitrarily large. Show that for any $p_0 < 1$, no matter how close to 1, there is a $p > p_0$ and a σ, such that the normal mixture with these values of p and σ has a kurtosis at least M. This shows that there is a normal mixture arbitrarily close to a normal distribution but with a kurtosis above any arbitrarily large value of M.

Proof:

From Problem 4(c), we know that $\kappa\left(Y|p,\sigma^2\right) \geq M$ when

$$\sigma^2 \geq \sigma_M^2 = \frac{2\sqrt{M^2+3M}+2M+3}{3} > 1 \quad \text{and} \quad p = \frac{\sigma^2}{1+\sigma^2}.$$

From the conclusion above, we set p_M as

$$p_M = \frac{\sigma_M^2}{1+\sigma_M^2}$$

If $p_0 < p_M$, we choose $p \geq p_M$, which satisfies

$$\sigma^2 = \frac{p}{1-p} \geq \frac{p_M}{1-p_M} = \sigma_M^2.$$

If $p_0 \geq p_M$, for any $p > p_0$, we have

$$\sigma^2 = \frac{p}{1-p} \geq \frac{p_0}{1-p_0} \geq \frac{p_M}{1-p_M} = \sigma_M^2.$$

Since we can always choose a $\sigma^2 \geq \sigma_M^2$, then

$$\kappa\big(Y|p, \sigma^2\big) \geq M.$$

Hence, we can always choose $p > \max\{p_M, p_0\}$ and $\sigma = \sqrt{\frac{p}{1-p}}$, such that the normal mixture with these values of p and σ has a kurtosis at least M, which shows that there is a normal mixture arbitrarily close to a normal distribution but with a kurtosis above any arbitrarily large value of M. ■

Problem 2

For any univariate parameter θ and its estimator $\hat{\theta}$, we define the bias to be $b\big(\hat{\theta}\big) = \mathbb{E}\big(\hat{\theta}\big) - \theta$ and the MSE to be $\mathrm{MSE}\big(\hat{\theta}\big) = \mathbb{E}\big(\hat{\theta} - \theta\big)^2$. Show that

$$\mathrm{MSE}\big(\hat{\theta}\big) = b^2\big(\hat{\theta}\big) + \mathrm{Var}\big(\hat{\theta}\big).$$

Proof:

Recall the definitions of $\mathrm{MSE}\big(\hat{\theta}\big)$, $b\big(\hat{\theta}\big)$ and $\mathrm{Var}\big(\hat{\theta}\big)$, we have

$$\begin{aligned}
\mathrm{MSE}\big(\hat{\theta}\big) &= \mathbb{E}\big(\hat{\theta} - \theta\big)^2 \\
&= \mathbb{E}\Big[\hat{\theta} - \mathbb{E}\big(\hat{\theta}\big) + \mathbb{E}\big(\hat{\theta}\big) - \theta\Big]^2 \\
&= \mathbb{E}\Big[\hat{\theta} - \mathbb{E}\big(\hat{\theta}\big)\Big]^2 + 2\Big\{\mathbb{E}\Big[\hat{\theta} - \mathbb{E}\big(\hat{\theta}\big)\Big]\Big\}\Big\{\mathbb{E}\Big[\mathbb{E}\big(\hat{\theta}\big) - \theta\Big]\Big\} + \mathbb{E}\Big[\mathbb{E}\big(\hat{\theta}\big) - \theta\Big]^2 \\
&= \mathrm{Var}\big(\hat{\theta}\big) + 2\Big[\mathbb{E}\big(\hat{\theta}\big) - \mathbb{E}\big(\hat{\theta}\big)\Big]\Big[\mathbb{E}\big(\hat{\theta}\big) - \theta\Big] + \Big(\mathbb{E}\big(\hat{\theta}\big) - \theta\Big)^2 \\
&= \mathrm{Var}\big(\hat{\theta}\big) + b^2\big(\hat{\theta}\big).
\end{aligned}$$

■

Problem 3

(a) Suppose $X_1, X_2, \cdots, X_n$ are i.i.d. $\mathcal{N}\big(\mu, \sigma^2\big)$ where $\big(\mu, \sigma^2\big)$ are both unknown. Find the MLE of $\big(\mu, \sigma^2\big)$.

Solution:

Recall the density function of normal distribution, we have

$$f_{X_i}\big(x_i \mid \mu, \sigma^2\big) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}, \; i = 1, 2, \cdots, n.$$

Then the joint distribution, which has the same form as likelihood function, can be written as

$$\begin{aligned}
L\big(\mu, \sigma^2 \mid x_1, x_2, \cdots, x_n\big) &= f_{X_1, X_2, \cdots, X_n}\big(x_1, x_2, \cdots, x_n \mid \mu, \sigma^2\big) \\
&= \prod_{i=1}^{n} f_{X_i}\big(x_i \mid \mu, \sigma^2\big) \\
&= \frac{1}{\big(2\pi\sigma^2\big)^{\frac{n}{2}}} \mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2},
\end{aligned}$$

and the log-likelihood function is

$$
\begin{aligned}
l\big(\mu, \sigma^2 \mid x_1, x_2, \cdots, x_n\big) &= \log\Big(L\big(\mu, \sigma^2 \mid x_1, x_2, \cdots, x_n\big)\Big) \\
&= \log\left(\frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}} \mathrm{e}^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}\right) \\
&= -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\mu\right)^2.
\end{aligned}
$$

Calculate the derivatives for μ and σ^2, respectively. Let them be zero and we get the MLE of μ and σ^2

$$
\begin{aligned}
\left.\frac{\partial l}{\partial \mu}\right|_{\sigma^2=\hat{\sigma}^2} &= -\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i-\mu\right) = 0, \\
\left.\frac{\partial l}{\partial \sigma^2}\right|_{\mu=\hat{\mu}} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left(x_i-\hat{\mu}\right)^2 = \frac{1}{2\sigma^2}\left(\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 - n\right) = 0, \\
&\Rightarrow\ \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}x_i = \bar{x},\ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 = \frac{n-1}{n}S^2,
\end{aligned}
$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$.

(b) Suppose $X_1, X_2, \cdots, X_n$ are i.i.d. $\mathrm{Exp}\big(\lambda\big)$ where λ is the unknown scale parameter. Find the MLE of λ.

Solution:

Recall the density function of exponential distribution, we have

$$
f_{X_i}\big(x_i \mid \lambda\big) = \lambda\mathrm{e}^{-\lambda x_i},\ i = 1, 2, \cdots, n.
$$

Then the joint distribution can be written as

$$
\begin{aligned}
L\big(\lambda \mid x_1, x_2, \cdots, x_n\big) &= f_{X_1, X_2, \cdots, X_n}\big(x_1, x_2, \cdots, x_n \mid \lambda\big) \\
&= \prod_{i=1}^{n} f_{X_i}\big(x_i \mid \mu, \sigma^2\big) \\
&= \lambda^n \mathrm{e}^{-\lambda\sum_{i=1}^{n}x_i},
\end{aligned}
$$

and the log-likelihood function is

$$
\begin{aligned}
l\big(\lambda \mid x_1, x_2, \cdots, x_n\big) &= \log\Big(L\big(\lambda \mid x_1, x_2, \cdots, x_n\big)\Big) \\
&= \log\left(\lambda^n \mathrm{e}^{-\lambda\sum_{i=1}^{n}x_i}\right) \\
&= n\log\big(\lambda\big) - \lambda\sum_{i=1}^{n}x_i.
\end{aligned}
$$

Calculate the derivatives for λ. Let it be zero and we get the MLE of λ

$$
\begin{aligned}
\frac{\partial l}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^{n}x_i = 0, \\
\Rightarrow\ \hat{\lambda} &= \frac{1}{n}\sum_{i=1}^{n}x_i = \bar{x},
\end{aligned}
$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i$.

Problem 4

Use the data in your Project 1 and assume the daily log returns are i.i.d. $\mathcal{N}(\mu, \sigma^2)$. Write a program to compute the MLE of (μ, σ^2) for each year.

Solution:

The data are chosen from stock dataset of Google Inc. from 2012-01-01 to 2018-12-31.

```
raw.data <- read.csv('../PJ1/data/GOOG.csv')
stock <- data.frame(date=raw.data$Date[-1], price=raw.data$Close[-1],
    log.return=log(raw.data$Close[2:(dim(raw.data)[1])]) -
        log(raw.data$Close[1:(dim(raw.data)[1] - 1)]))
NegLogLikeliFuncNorm <- function (para, x) {
    return(-sum(log(dnorm(x, mean=para[1], sd=sqrt(para[2])))))
}

result <- data.frame()
ratio <- 10
for (year in 2012:2018) {
    stock.year <- stock[grep(as.character(year), stock$date), ]
    para <- optim(par=c(0, 0.005), fn=NegLogLikeliFuncNorm,
        x=ratio * stock.year$log.return, hessian=T)
    result <- rbind(result, c(year, dim(stock.year)[1],
        format(para$par[1] / ratio, scientific=F, digits=4, nsmall=4),
        format(sqrt(solve(para$hessian)[1, 1]) / ratio,
               scientific=F, digits=4, nsmall=4),
        format(para$par[2] / (ratio^2), scientific=F, digits=4, nsmall=4),
        format(sqrt(solve(para$hessian)[2, 2]) / (ratio^2),
               scientific=F, digits=4, nsmall=4)))
}

knitr::kable(
    x=result, booktabs=T,
    caption='Summary for daily log returns (normal distribution)',
    col.names=c('Year', '# of Obs.',
        '$\\hat{\\mu}$', 'se($\\hat{\\mu}$)',
        '$\\hat{\\sigma}^2$', 'se($\\hat{\\sigma}^2$)'),
    align=c('c', 'c', 'c', 'c', 'c', 'c'))
```

Table 1: Summary for daily log returns (normal distribution)

Year	# of Obs.	$\hat{\mu}$	se($\hat{\mu}$)	$\hat{\sigma}^2$	se($\hat{\sigma}^2$)
2012	249	0.0002449	0.000917	0.0002094	0.00001864
2013	252	0.001826	0.0008439	0.0001795	0.00001584
2014	252	-0.0002444	0.0008419	0.0001786	0.00001576
2015	252	0.001464	0.001138	0.0003265	0.00002901
2016	252	0.00006679	0.0007943	0.000159	0.00001399
2017	251	0.001212	0.0006129	0.00009428	0.000008134
2018	250	-0.00004063	0.001125	0.0003165	0.00002823

The table above is a summary of the MLEs of daily log returns. From the results, we can find that the MLEs of the means and variances are almost zero. Hence, we can conclude that the log returns are located around zero and fluctuate around zero in relatively small ranges, which means $\mathcal{N}(\mu, \sigma^2)$ is not a suitable model.

Problem 5*

Do the same anaylsis in Problem 3 but assume $t_\nu(\mu, \lambda^2)$ for the daily log returns where ν is the degreee of freedom, μ, λ are the location and scale parameters respectively. Compare your results in Problem 3 and 4.

Solution:

Recall the density function of t location-scale distribution $t_\nu(\mu, \lambda^2)$, we have

$$f_{X_i}(x_i \mid \nu, \mu, \lambda^2) = \frac{C(\nu)}{\lambda}\left(1 + \frac{(x_i - \mu)^2}{\nu\lambda^2}\right)^{-\frac{\nu+1}{2}}, \ i = 1, 2, \cdots, n.$$

where $C(\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{(\pi\nu)^{\frac{1}{2}}\Gamma(\frac{\nu}{2})}$. Then the joint distribution, which has the same form as likelihood function, can be written as

$$\begin{aligned}
L(\nu, \mu, \lambda^2 \mid x_1, x_2, \cdots, x_n) &= f_{X_1, X_2, \cdots, X_n}(x_1, x_2, \cdots, x_n \mid \nu, \mu, \lambda^2) \\
&= \prod_{i=1}^{n} f_{X_i}(x_i \mid \nu, \mu, \lambda^2) \\
&= \frac{C^n(\nu)}{\lambda^n}\left\{\prod_{i=1}^{n}\left(1 + \frac{(x_i - \mu)^2}{\nu\lambda^2}\right)\right\}^{-\frac{\nu+1}{2}},
\end{aligned}$$

and the log-likelihood function is

$$\begin{aligned}
l(\nu, \mu, \lambda^2 \mid x_1, x_2, \cdots, x_n) &= \log\Big(L(\nu, \mu, \lambda^2 \mid x_1, x_2, \cdots, x_n)\Big) \\
&= \log\left(\frac{C^n(\nu)}{\lambda^n}\left\{\prod_{i=1}^{n}\left(1 + \frac{(x_i - \mu)^2}{\nu\lambda^2}\right)\right\}^{-\frac{\nu+1}{2}}\right) \\
&= n\log C(\nu) - \frac{n}{2}\log(\lambda^2) - \frac{\nu+1}{2}\sum_{i=1}^{n}\log\left(1 + \frac{(x_i - \mu)^2}{\nu\lambda^2}\right).
\end{aligned}$$

Calculate the derivatives for ν, μ and λ^2, respectively, and let them be zero

$$\left.\frac{\partial l}{\partial \nu}\right|_{\mu=\hat{\mu},\ \lambda^2=\hat{\lambda}^2} = \frac{nC'(\nu)}{C(\nu)} - \frac{1}{2}\sum_{i=1}^{n}\log\left(1 + \frac{(x_i - \mu)^2}{\nu\lambda^2}\right) + \frac{\nu+1}{2\nu}\sum_{i=1}^{n}\frac{(x_i - \mu)^2}{\nu\lambda^2 + (x_i - \mu)^2} = 0,$$

$$\left.\frac{\partial l}{\partial \mu}\right|_{\nu=\hat{\nu},\ \lambda^2=\hat{\lambda}^2} = (\nu+1)\sum_{i=1}^{n}\frac{x_i - \mu}{\nu\lambda^2 + (x_i - \mu)^2} = 0,$$

$$\left.\frac{\partial l}{\partial \lambda^2}\right|_{\nu=\hat{\nu},\ \mu=\hat{\mu}} = -\frac{n}{2\lambda^2} + \frac{\nu+1}{2\lambda^2}\sum_{i=1}^{n}\frac{(x_i - \mu)^2}{\nu\lambda^2 + (x_i - \mu)^2} = 0,$$

Unfortunately, we can't derive an explicit expression for the MLE of $t_\nu(\mu, \lambda^2)$. Hence, we should use numerical methods to find the MLE of ν, μ and λ^2 for $t_\nu(\mu, \lambda^2)$ distribution.

```
raw.data <- read.csv('../PJ1/data/GOOG.csv')
stock <- data.frame(date=raw.data$Date[-1], price=raw.data$Close[-1],
    log.return=log(raw.data$Close[2:(dim(raw.data)[1])]) -
        log(raw.data$Close[1:(dim(raw.data)[1] - 1)]))
NegLogLikeliFuncT  <- function (para, x) {
    return(-sum(log(dt((x - para[1]) / sqrt(para[2]), df=para[3]) / sqrt(para[2]))))
```

```r
}

result <- data.frame()
ratio <- 10
for (year in 2012:2018) {
    stock.year <- stock[grep(as.character(year), stock$date), ]
    para <- optim(par=c(0, 0.0001, 1), fn=NegLogLikeliFuncT,
        x=ratio * stock.year$log.return, hessian=T)
    result <- rbind(result, c(year, dim(stock.year)[1],
        format(para$par[1] / ratio, scientific=F, digits=4, nsmall=4),
        format(sqrt(solve(para$hessian)[1, 1]) / (ratio^2),
               scientific=F, digits=4, nsmall=4),
        format(para$par[2] / (ratio^2), scientific=F, digits=4, nsmall=4),
        format(sqrt(solve(para$hessian)[2, 2]) / (ratio^2),
               scientific=F, digits=4, nsmall=4),
        format(para$par[3], scientific=F, digits=4, nsmall=4),
        format(sqrt(solve(para$hessian)[3, 3]),
               scientific=F, digits=4, nsmall=4)))
}

knitr::kable(
    x=result, booktabs=T,
    caption='Summary for daily log returns (t distribution)',
    col.names=c('Year', '# of Obs.',
        '$\\hat{\\mu}$', 'se($\\hat{\\mu}$)',
        '$\\hat{\\lambda}^2$', 'se($\\hat{\\lambda}^2$)',
        '$\\hat{\\nu}$', 'se($\\hat{\\nu}$)'),
    align=c('c', 'c', 'c', 'c', 'c', 'c', 'c'))
```

Table 2: Summary for daily log returns (t distribution)

Year	# of Obs.	$\hat{\mu}$	se($\hat{\mu}$)	$\hat{\lambda}^2$	se($\hat{\lambda}^2$)	$\hat{\nu}$	se($\hat{\nu}$)
2012	249	0.0009561	0.00007511	0.00009999	0.00001429	4.0808	0.9879
2013	252	0.0006348	0.00006463	0.00007348	0.00001003	4.2633	1.0409
2014	252	0.00009945	0.00007988	0.0001265	0.00001916	6.6250	2.7322
2015	252	0.0001686	0.00008492	0.000124	0.00001876	3.4609	0.7662
2016	252	0.0007687	0.00006663	0.0000781	0.00001272	3.5603	0.9066
2017	251	0.001406	0.00005128	0.00004624	0.000006447	3.6535	0.8676
2018	250	0.001141	0.0001006	0.0001761	0.00003073	4.0271	1.2064

Compare the results between Problem 4 and 5, we can find that the parameters are similar. The MLEs of μ and λ^2 are almost zero. The MLEs of ν are between 3 and 6, which means $t_\nu(\mu, \lambda^2)$ have heavier tails than normal distribution. But similarly, $t_\nu(\mu, \lambda^2)$ is also not a suitable model for log returns.

Problem 6*

(a) Show that for any non-negative random variable X,

$$\mathbb{E}(X) = \int_0^{+\infty} P(X > t)\ \mathrm{d}t.$$

Proof:

Recall the relationship between density function and probability, we have

$$P\big(X > t\big) = \int_t^{+\infty} f_X\big(x\big)\ \mathrm{d}x.$$

Hence, consider the following integration and we get

$$\begin{aligned}
\int_0^{+\infty} P\big(X > t\big)\ \mathrm{d}t &= \int_0^{+\infty} \left(\int_t^{+\infty} f_X\big(x\big)\ \mathrm{d}x\right)\ \mathrm{d}t \\
&= \int_0^{+\infty} \left(\int_0^{x} f_X\big(x\big)\ \mathrm{d}t\right)\ \mathrm{d}x \\
&= \int_0^{+\infty} \left(t f_X\big(x\big)\Big|_0^x\right)\ \mathrm{d}x \\
&= \int_0^{+\infty} x f_X\big(x\big)\ \mathrm{d}x \\
&= \mathbb{E}\big(X\big).
\end{aligned}$$

∎

(b) Show that for any random variable X,

$$\mathbb{E}\big(|X|^p\big) = p\int_0^{+\infty} t^{p-1} P\big(|X| > t\big)\ \mathrm{d}t.$$

Proof:

Similarly, we have the following relationship

$$P\big(|X| > t\big) = \int_{-\infty}^{-t} f_X(x)\ \mathrm{d}x + \int_t^{+\infty} f_X(x)\ \mathrm{d}x.$$

Hence, consider the following integration and we get

$$\begin{aligned}
p\int_0^{+\infty} t^{p-1} P\big(|X| > t\big)\ \mathrm{d}t &= p\int_0^{+\infty} t^{p-1}\left(\int_{-\infty}^{-t} f_X(x)\ \mathrm{d}x + \int_t^{+\infty} f_X(x)\ \mathrm{d}x\right)\ \mathrm{d}t \\
&= \int_0^{+\infty}\left(\int_{-\infty}^{-t} pt^{p-1} f_X(x)\ \mathrm{d}x\right)\ \mathrm{d}t + \int_0^{+\infty}\left(\int_t^{+\infty} pt^{p-1} f_X(x)\ \mathrm{d}x\right)\ \mathrm{d}t \\
&= \int_{-\infty}^{0}\left(\int_0^{-x} pt^{p-1} f_X(x)\ \mathrm{d}t\right)\ \mathrm{d}x + \int_0^{+\infty}\left(\int_0^{x} pt^{p-1} f_X(x)\ \mathrm{d}t\right)\ \mathrm{d}x \\
&= \int_{-\infty}^{0}\left(t^p f_X(x)\Big|_0^{-x}\right)\ \mathrm{d}x + \int_0^{+\infty}\left(t^p f_X(x)\Big|_0^{x}\right)\ \mathrm{d}x \\
&= \int_{-\infty}^{0} |x|^p f_X(x)\ \mathrm{d}x + \int_0^{+\infty} |x|^p f_X(x)\ \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} |x|^p f_X(x)\ \mathrm{d}x \\
&= \mathbb{E}\big(|X|^p\big).
\end{aligned}$$

∎