Homework 5 Solutions

46-923, Fall 2017

December 9, 2017

You can submit separate pdf files, one generated from the R Markdown, and the other from the "derivations" required in Questions 2 and 3. The relevant .Rmd file should also be submitted.

Please do not submit photos of your homework. Scanners are available for your use.

Question 1

Redo the analysis on the Kellogg's data that we did during lecture, but this time look at the **weekly** log returns instead of the daily log returns. Show all of the R code needed to perform the analysis.

1. Reconstruct the plot comparing the five fitted models that we considered during lecture: normal, GED, nonstandard T, skewed T, and skewed T with shift.

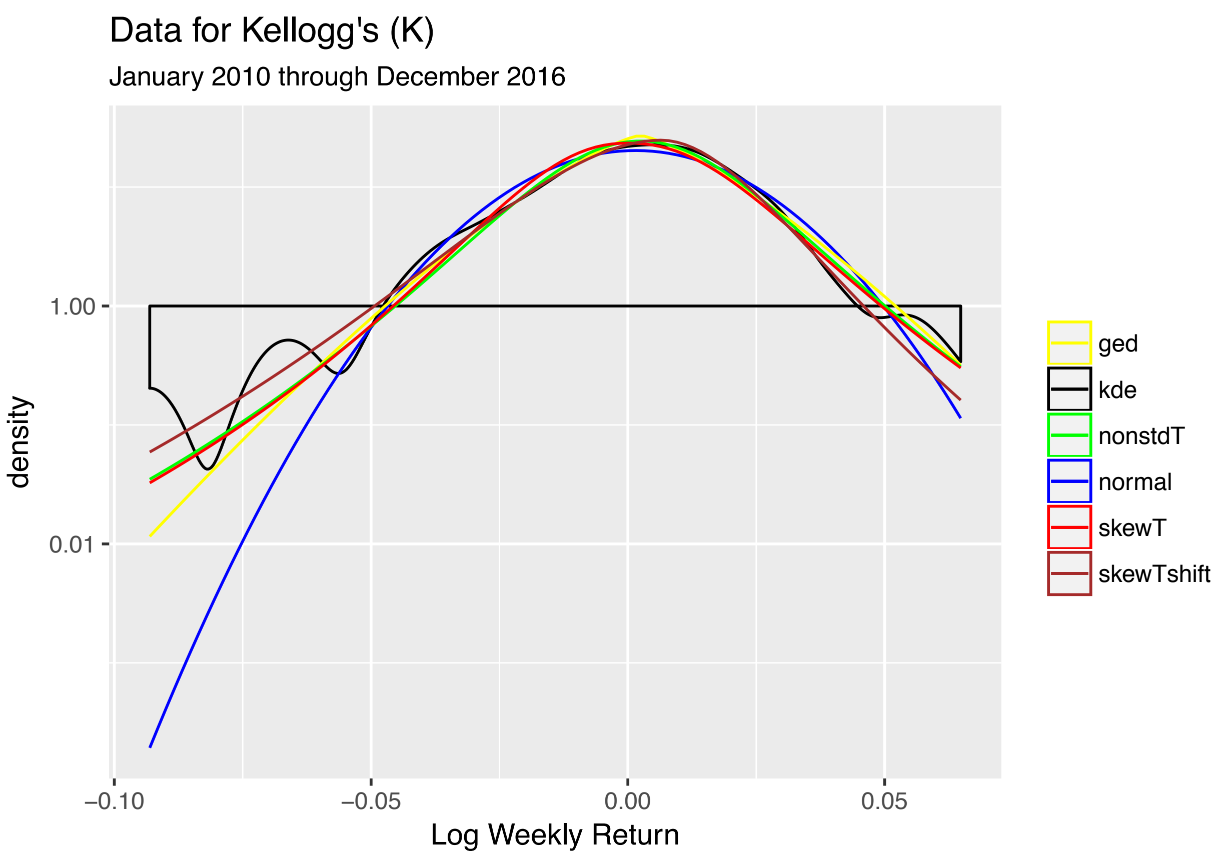
```r
library(ggplot2)
library(fGarch)
source("http://www.stat.cmu.edu/~cschafer/MSCF/ModelSelectionExample.txt")

# Downloading data
Kellogg = getSymbols("K",from="2010-1-1",to="2016-12-31", auto.assign=F)
ldrK = data.frame(weeklyReturn(Ad(Kellogg), type="log"))

# Fitting Models
n = length(ldrK$weekly.returns)
normalout = c(mean(ldrK$weekly.returns),
              sqrt(var(ldrK$weekly.returns)*(n-1)/n))
gedout = FitGED(ldrK$weekly.returns)
nonstdTout = FitNonStdT(ldrK$weekly.returns)
skewtout = FitSkewT(ldrK$weekly.returns)
skewtshiftout = FitSkewT(ldrK$weekly.returns,allowshift=T,
                         control=list(maxit=1000))

# Setting up Plot
ggplot(ldrK,aes(x=weekly.returns)) +
  geom_density(bw="SJ",aes(color="kde")) +
  stat_function(fun=dnorm, aes(color="normal"),
                args=list(mean=normalout[1],
                          sd=normalout[2])) +
  stat_function(fun=dged, aes(color="ged"),
                args=list(mean=gedout$mle[1],sd=gedout$mle[2],
                          nu=gedout$mle[3])) +
```

```r
stat_function(fun=dstd, aes(color="nonstdT"),
              args=list(mean=nonstdTout$mle[1],sd=nonstdTout$mle[2],
                        nu=nonstdTout$mle[3])) +
stat_function(fun=dSkewT, aes(color="skewT"),
              args=list(k=skewtout$mle[1],n=skewtout$mle[2],
                        lambda=skewtout$mle[3],
                        sigma2=skewtout$mle[4])) +
stat_function(fun=dSkewT, aes(color="skewTshift"),
              args=list(k=skewtshiftout$mle[1],n=skewtshiftout$mle[2],
                        lambda=skewtshiftout$mle[3],
                        sigma2=skewtshiftout$mle[4],
                        shift=skewtshiftout$mle[5])) +
scale_color_manual(name="",values=c("kde"="black","normal"="blue",
                                    "ged"="yellow","nonstdT"="green",
                                    "skewT"="red",
                                    "skewTshift"="brown")) +
labs(x="Log Weekly Return",title="Data for Kellogg's (K)",
     subtitle="January 2010 through December 2016") +
scale_y_log10()
```



2. Which of the five models is preferred by AIC?

```r
# Maximum likelihood for all of the models
gedML = sum(dged(ldrK$weekly.returns,gedout$mle[1],gedout$mle[2],
                 gedout$mle[3],log=T))
nonStandardT_ML = sum(dstd(ldrK$weekly.returns,nonstdTout$mle[1],
                           nonstdTout$mle[2],nonstdTout$mle[3],log=T))
skewT_ML = sum(dSkewT(ldrK$weekly.returns,skewtout$mle[1],skewtout$mle[2],
                      skewtout$mle[3],skewtout$mle[4],log=T))
skewTshiftML = sum(dSkewT(ldrK$weekly.returns,skewtshiftout$mle[1],
                          skewtshiftout$mle[2],skewtshiftout$mle[3],
                          skewtshiftout$mle[4],skewtshiftout$mle[5],log=T))
normML = sum(dnorm(ldrK$weekly.returns,normalout[1],normalout[2],log=T))

# Now calculate each AIC of each
aic_func = function (ML, params) {
  return (-2*ML + 2*params)
}

aic_func(gedML, 3)
```

```
## [1] -1846.876
```

```r
aic_func(nonStandardT_ML, 3)
```

```
## [1] -1850.368
```

```r
aic_func(skewT_ML, 4)
```

```
## [1] -1843.426
```

```r
aic_func(skewTshiftML, 5)
```

```
## [1] -1851.83
```

```r
aic_func(normML, 2)
```

```
## [1] -1828.087
```

Smallest AIC corresponds to the skewed T with shift - even though they are all pretty close.

46-923, Financial Data Science II
Homework 5 Solutions

Carnegie Mellon University

December 9, 2017

Question 2

Consider the example in lecture where we estimated the pair (α, σ^2) from the geometric Brownian motion. Suppose that instead I wanted to estimate (α, σ), i.e., I want to determine the asymptotic distribution of the MLE for this parameter vector. What is the asymptotic distribution for the MLE $(\hat{\alpha}, \hat{\sigma})$?

Following the notes of Lecture 5 (November 29), we have that considering $Y_1, Y_2, ..., Y_n$ i.i.d. samples from the geometric Brownian motion $Y_i \sim N(\Delta\mu, \Delta\sigma^2)$ $\forall i = 1, ..., n$.
Let $\theta_1 = \Delta\mu$ and $\theta_2 = \Delta\sigma^2$ the original parameters and let $\Theta_0 = (\theta_1, \theta_2)$.
Derivations in class showed that $\hat{\Theta}_{MLE} = (\hat{\theta}_1, \hat{\theta}_2) = (\overline{Y}, \frac{n-1}{n} S_y^2)$ has an asymptotic distribution which is $N(\Theta_0, \frac{\mathbb{I}^{-1}(\Theta_0)}{n})$, where the covariance matrix is equal to:

$$\frac{\mathbb{I}^{-1}(\Theta_0)}{n} = \begin{bmatrix} \frac{\theta_2}{n} & 0 \\ 0 & \frac{2\theta_2^2}{n} \end{bmatrix} \tag{1}$$

In our case we want to identify the asymptotic distribution for $(\hat{\alpha}, \hat{\sigma})$.
Using the same notation as the multivariate Delta method we have that:

$$\alpha = g_1(\Theta_0) = \mu + \frac{\sigma^2}{2} = \frac{\theta_1}{\Delta} + \frac{\theta_2}{2\Delta} \tag{2}$$

$$\sigma = g_2(\Theta_0) = \sqrt{\frac{\theta_2}{\Delta}} \tag{3}$$

By invariance of the MLE we have that:

$$\hat{\alpha}_{MLE} = g_1(\Theta_{MLE}) = \frac{\hat{\theta}_1}{\Delta} + \frac{\hat{\theta}_2}{2\Delta} \tag{4}$$

$$\hat{\sigma}_{MLE} = g_2(\Theta_{MLE}) = \sqrt{\frac{\hat{\theta}_2}{\Delta}} \tag{5}$$

Using the multivariate CLT we have that the asymptotic distribution is a bivariate normal.

The mean for $(\hat{\alpha}, \hat{\sigma})$ is (α, σ), which is a consequence of the unbiasedness of $\hat{\theta_1}$ for and $\hat{\theta_2}$ and continuous mapping theorem[1].
In order to calculate the covariance matrix we apply the multivariate delta method, first by computing the matrix $\tilde{G}$:

$$\tilde{G} = \begin{bmatrix} \frac{\partial g_1(\Theta_0)}{\partial \theta_1} & \frac{\partial g_1(\Theta_0)}{\partial \theta_2} \\ \frac{\partial g_2(\Theta_0)}{\partial \theta_1} & \frac{\partial g_2(\Theta_0)}{\partial \theta_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\Delta} & \frac{1}{2\Delta} \\ 0 & \frac{1}{2}\left(\frac{\theta_2}{\Delta}\right)^{-1/2} \frac{1}{\Delta} \end{bmatrix} = \begin{bmatrix} \frac{1}{\Delta} & \frac{1}{2\Delta} \\ 0 & \frac{1}{2\sigma\Delta} \end{bmatrix} \tag{6}$$

And then considering the product $\tilde{G}\Sigma\tilde{G}^T$:

$$\tilde{G}\Sigma\tilde{G}^T = \tilde{G}\frac{\mathbb{I}^{-1}(\Theta_0)}{n}\tilde{G}^T = \begin{bmatrix} \frac{1}{\Delta} & \frac{1}{2\Delta} \\ 0 & \frac{1}{2\sigma\Delta} \end{bmatrix} \begin{bmatrix} \frac{\Delta\sigma^2}{n} & 0 \\ 0 & \frac{2\Delta^2\sigma^4}{n} \end{bmatrix} \begin{bmatrix} \frac{1}{\Delta} & 0 \\ \frac{1}{2\Delta} & \frac{1}{2\sigma\Delta} \end{bmatrix} \tag{7}$$

$$= \begin{bmatrix} \frac{\sigma^3}{\Delta n} + \frac{\sigma^4}{2n} & \frac{\sigma^3}{2n} \\ \frac{\sigma^3}{2n} & \frac{\sigma^2}{2n} \end{bmatrix} = \tilde{\Sigma} \tag{8}$$

Hence the asymptotic distribution of $(\hat{\alpha}, \hat{\sigma})$ is $N((\alpha, \sigma), \tilde{\Sigma})$.

Question 3

Suppose that $X_1, X_2, ..., X_n$ are i.i.d. random variables. Each X_i takes three possible values (call these 1, 2, and 3, but the names are unimportant). The random variables are such that:

$$\mathbb{P}(X_1 = 1) = p_1 \quad , \quad \mathbb{P}(X_2 = 1) = p_2 \quad , \quad \mathbb{P}(X_3 = 1) = 1 - p_1 - p_2 \tag{9}$$

The following restrictions are placed on the parameters:

$$p_1 > 0 \quad , \quad p_2 > 0 \quad , \quad p_1 + p_2 \leq 1 \tag{10}$$

1. Derive the MLE for $\theta = (p_1, p_2)$.

As per hint, let's define n_j the number of the X_i which are equal to j , $j = 1, 2, 3$. On top $n = \sum_{j=1}^{3} n_j$, the likelihood function can be then written as:

$$\mathbb{L}(\theta) = p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n_3} \tag{11}$$

This implies that the log-likelihood is as follows:

$$\log \mathbb{L}(\theta) = n_1 \log(p_1) + n_2 \log(p_2) + n_3 \log(1 - p_1 - p_2) \tag{12}$$

Solving therefore the following system we have that:

$$\begin{cases} \frac{\partial \log \mathbb{L}(\theta)}{\partial p_1} = 0 \\ \frac{\partial \log \mathbb{L}(\theta)}{\partial p_2} = 0 \end{cases} \implies \begin{cases} p_1 = \frac{n_1 - n_1 p_2}{n_3 + n_1} \\ p_2 = \frac{n_2 - n_2 p_1}{n_3 + n_2} \end{cases} \implies \hat{p}_i = \frac{n_i}{n}, \quad i = 1, 2 \tag{13}$$

[1] Since $\frac{\theta_2}{\Delta} \xrightarrow{P} \sigma^2$ then $\sqrt{\frac{\hat{\theta_2}}{\Delta}} \xrightarrow{P} \sigma$.

2. What is the asymptotic distribution of the MLE $(\hat{p_1}, \hat{p_2})$?

By CLT the multivariate normal we have that the asymptotic distribution of $(\hat{p_1}, \hat{p_2})$ is bivariate normal with mean (p_1, p_2) and covariance matrix equal to $\frac{\mathbb{I}^{-1}(\theta)}{n}$. We need to calculate such covariance matrix.
The first step is calculating the Fisher Information matrix:

$$n\mathbb{I}(\theta) = \begin{bmatrix} \frac{\partial^2 \log \mathbb{L}(\theta)}{\partial p_1^2} & \frac{\partial^2 \log \mathbb{L}(\theta)}{\partial p_1 p_2} \\ \frac{\partial^2 \log \mathbb{L}(\theta)}{\partial p_1 p_2} & \frac{\partial^2 \log \mathbb{L}(\theta)}{\partial p_2^2} \end{bmatrix} = \begin{bmatrix} -\frac{n_1}{p_1^2} - \frac{n_3}{(1-p_1-p_2)^2} & -\frac{n_3}{(1-p_1-p_2)^2} \\ -\frac{n_3}{(1-p_1-p_2)^2} & -\frac{n_2}{p_2^2} - \frac{n_3}{(1-p_1-p_2)^2} \end{bmatrix} \tag{14}$$

$$= n \begin{bmatrix} \frac{1}{p_1} + \frac{1}{1-p_1-p_2} & \frac{1}{1-p_1-p_2} \\ \frac{1}{1-p_1-p_2} & \frac{1}{p_2} + \frac{1}{1-p_1-p_2} \end{bmatrix} \tag{15}$$

And then we invert it obtaining, after calculating that $\det(\mathbb{I}(\theta)) = (1 - p_1 - p_2)p_1p_2$:

$$\frac{\mathbb{I}^{-1}(\theta)}{n} = \frac{(1-p_1-p_2)p_1p_2}{n} \begin{bmatrix} \frac{1}{p_2} + \frac{1}{1-p_1-p_2} & -\frac{1}{1-p_1-p_2} \\ -\frac{1}{1-p_1-p_2} & \frac{1}{p_1} + \frac{1}{1-p_1-p_2} \end{bmatrix} \tag{16}$$

$$= \begin{bmatrix} \frac{(1-p_1)p_1}{n} & -\frac{p_1p_2}{n} \\ -\frac{p_1p_2}{n} & \frac{(1-p_2)p_2}{n} \end{bmatrix} = \Sigma_{12} \tag{17}$$

Extra Practice

Derive a $100(1-\alpha)\%$ confidence interval for p_1/p_2.

This is again an application of the Delta Method.
If we consider the function as $g(\theta)$ we can write it down as:

$$g(\theta) = \frac{p_1}{p_2} \tag{18}$$

For the invariance of the MLE then $\hat{r}_{MLE} = \widehat{\left(\frac{p_1}{p_2}\right)}_{MLE} = \frac{n_1}{n_2}$.
The asymptotic distribution will be normal with mean $p1/p2$ and variance $\tilde{\Sigma}$.
In order to determine the variance we first obtain the gradient $\nabla_g(\theta)$ which is as follows:

$$\nabla_g(\theta) = \left(\frac{\partial g(\theta)}{\partial p_1}, \frac{\partial g(\theta)}{\partial p_2}\right) = \left(\frac{1}{p_2}, -\frac{p_1}{p_2^2}\right) \tag{19}$$

And then we obtain the covariance matrix by considering the covariance matrix Σ_{12} in equation 17 obtained for the multivariate distribution of $(\hat{p_1}, \hat{p_2})$:

$$\tilde{\Sigma} = \nabla_g(\theta)\Sigma_{12}\nabla_g(\theta)^T = \begin{bmatrix} \frac{1}{p_2} & -\frac{p_1}{p_2^2} \end{bmatrix} \begin{bmatrix} \frac{(1-p_1)p_1}{n} & -\frac{p_1p_2}{n} \\ -\frac{p_1p_2}{n} & \frac{(1-p_2)p_2}{n} \end{bmatrix} \begin{bmatrix} \frac{1}{p_2} \\ -\frac{p_1}{p_2^2} \end{bmatrix} \tag{20}$$

$$= \begin{bmatrix} \frac{1}{p_2 n} & -\frac{p_1}{p_2 n} \end{bmatrix} \begin{bmatrix} \frac{1}{p_2} \\ -\frac{p_1}{p_2^2} \end{bmatrix} \tag{21}$$

$$= \frac{p_1^2 + p_2}{p_2^3 n} \tag{22}$$

And hence the $100(1-\alpha)\%$ confidence interval for p_1/p_2 is the following, given $Z_{1-\frac{\alpha}{2}}$ the $1-\frac{\alpha}{2}$ percentile of a normal distribution:

$$\left[\frac{\hat{p_1}}{\hat{p_2}} - Z_{1-\frac{\alpha}{2}}\frac{1}{\sqrt{n}}\sqrt{\frac{\hat{p_1}^2+\hat{p_2}}{\hat{p_2}^3}}, \frac{\hat{p_1}}{\hat{p_2}} + Z_{1-\frac{\alpha}{2}}\frac{1}{\sqrt{n}}\sqrt{\frac{\hat{p_1}^2+\hat{p_2}}{\hat{p_2}^3}}\right] \tag{23}$$