Mgmt 237E: Empirical Methods in Finance
Homework 1: Solution

Prof. Lars A. Lochstoer

TAs: Chady Gemayel and Mahyar Kargar

January 13, 2019

Problem 1: Modeling heavy-tails with jumps

The distribution of stock returns has fat tails (see lecture 1). As a result, we need models that deliver fat-tailed distributions (e.g., to price options on stocks). One way of doing this is to introduce jumps in the model. This problem describes the building blocks of how to do that.

Consider a Bernoulli-distributed variable:

$$B_t = \begin{cases} 0 & \text{with probability } 1-p \\ 1 & \text{with probability } p \end{cases},$$

and two independent Standard Normal variables ε_t and δ_t. Define the jump as

$$J_t = B_t \left(\mu_J + \sigma_J \delta_t\right)$$

and let return follow the process:

$$r_t = \mu + \sigma \varepsilon_t + J_t.$$

1. Derive the mean, variance, skewness and excess kurtosis of this distribution.

   Suggested solution: We first calculate the cumulant generating function for the Bernoulli-normal mixture. We know that

$$k(\mu_j + \sigma_J\delta_t, s) = \log \mathbb{E} e^{s\mu_j + s\sigma_J\delta_t} = (s\mu_j + \frac{1}{2}s^2\sigma_j^2) \tag{1}$$

$$k(J_t, s) = \log[e^{k(\mu_j+\sigma_J\delta_t,s)}p + e^{k(0,s)}(1-p)] \tag{2}$$

$$= \log[\exp(s\mu_j + \frac{1}{2}s^2\sigma_j^2)p + 1 - p] \tag{3}$$

Then we can get the cumulant generating function for r_t as

$$k(r_t, s) = s\mu + \frac{1}{2}s^2\sigma^2 + \log[\exp(s\mu_j + \frac{1}{2}s^2\sigma_j^2)p + 1 - p] \tag{4}$$

Thus we can take derivatives to get moments:

$$\frac{\partial k}{\partial s} = \mu + \sigma^2 s + \frac{1}{\exp(s\mu_j + \frac{1}{2}s^2\sigma_j^2)p + 1 - p}\exp(s\mu_j + \frac{1}{2}s^2\sigma_j^2)p(\mu_j + \sigma_j^2 s) \tag{5}$$

Setting s=0, we have that $\kappa_1 = \mu + p\mu_j$. Now, we let $\exp(s\mu_j + \frac{1}{2}s^2\sigma_j^2) = A$ and that $\frac{\partial A}{\partial s} = \mu_j + s\sigma_j^2 = B$. We then have $\frac{\partial B}{\partial s} = \sigma_j^2$. Thus

$$\frac{\partial^2 k}{\partial s^2} = \sigma^2 + p\frac{\sigma_j^2 A^2 p + (1-p)B^2 + (1-p)\sigma_j^2 A}{(Ap + 1 - p)^2} \tag{6}$$

Setting s=0, we have that $\kappa_2 = \sigma^2 + p(1-p)\mu_j^2 + p\sigma_j^2$.

$$\frac{\partial^3 k}{\partial s^3} = p\frac{3p(1-p)\sigma_j^2 AB + 3(1-p)^2\sigma_j^2 B - 2p(1-p)B^3}{(Ap + 1 - p)^3} \tag{7}$$

Setting s=0, we have that $\kappa_3 = 3p(1-p)\sigma_j^2\mu_j - 2p^2(1-p)\mu_j^3$.

$$\frac{\partial^4 k}{\partial s^4} = p\frac{\begin{aligned}-12p^2(1-p)\sigma_j^2 AB^2 + 3p^2(1-p)\sigma_j^4 A^2 + 6p(1-p)^2\sigma_j^4 A\\ - 12p(1-p)^2\sigma_j^2 B^2 + 3(1-p)^3\sigma_j^4 + 6p^2(1-p)B^4\end{aligned}}{(Ap + 1 - p)^4} \tag{8}$$

Setting s=0, we have that $\kappa_4 = -12p^2(1-p)\sigma_j^2\mu_j^2 + 3p(1-p)\sigma_j^4 + 6p^3(1-p)\mu_j^4$.

Then we use the definition to generate mean, variance, skewness and kurtosis:

$$\text{mean} = \kappa_1 = \mu + p\mu_j \tag{9}$$

$$\text{variance} = \kappa_2 = \sigma^2 + p(1-p)\mu_j^2 + p\sigma_j^2 \tag{10}$$

$$\text{skewness} = \frac{\kappa_3}{\kappa_2^{3/2}} = \frac{3p(1-p)\sigma_j^2\mu_j - 2p^2(1-p)\mu_j^3}{(\sigma^2 + p(1-p)\mu_j^2 + p\sigma_j^2)^{3/2}} \tag{11}$$

$$\text{kurtosis} = \frac{\kappa_4}{\kappa_2^2} = \frac{-12p^2(1-p)\sigma_j^2\mu_j^2 + 3p(1-p)\sigma_j^4 + 6p^3(1-p)\mu_j^4}{(\sigma^2 + p(1-p)\mu_j^2 + p\sigma_j^2)^2} \tag{12}$$

2. Explain carefully why this Bernoulli-normal mixture is potentially a better model of returns than the log-normal model.

   Suggested solution: As we can observe from the graph of the daily or month log return, there are a lot of extreme values that fall off 4 standard derivation range if we assume they are normal. Adding jumps to the model certainly help capture these values better.

3. Suppose that log returns r_t are simply i.i.d. distributed $N(0.008, 0.063^2)$. Plot a simulated series of 600 observations (50 years). Does it look like the data? What is missing?

   Next, suppose that log returns r_t are given by the above jump model with the following 5 parameters:

   $$(\mu, \sigma, p, \mu_J, \sigma_J) = (0.012, 0.05, 0.15, -0.03, 0.1)$$

   Using these parameter estimates, what are the unconditional mean, standard deviation, skewness and kurtosis of log stock returns? Again, plot a simulated series of 600 observations (50 years). Does it look like the data? What is missing?

   Suggested solution: Figure 1 shows the time series log return of the log normal sample we generated. It does not look like the real data in the sense that the return is pretty centered with no extreme value falls into the tail. Figure 2 presents the time series log return of the sample that we generated from the jump model. Compared to the simple log normal case, this jump model looks more like the real data in the sense that it does have a fatter tail and is able to generate some of the extreme events. However, it is still not quite as extreme as it is in the data. Table 1 shows the sample moments for the sample generated from the jump model. Though it seems the model

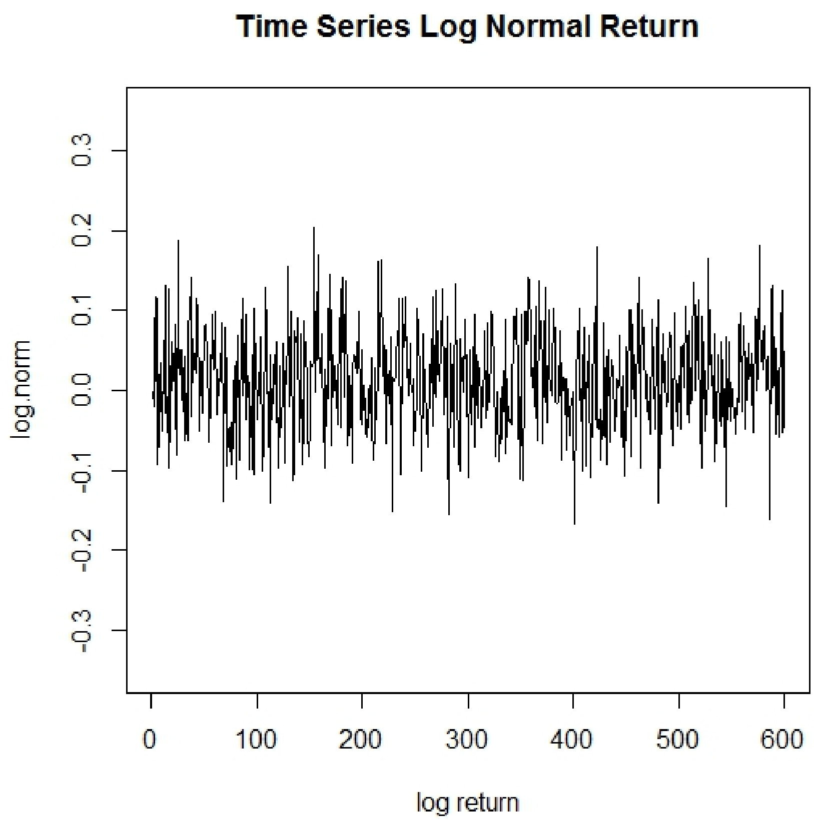

Figure 1: Time Series Log Return of the Log-normal Model

is still lack of volatility clustering, the jump model is approaching the real data with a negative skewness and excess kurtosis.

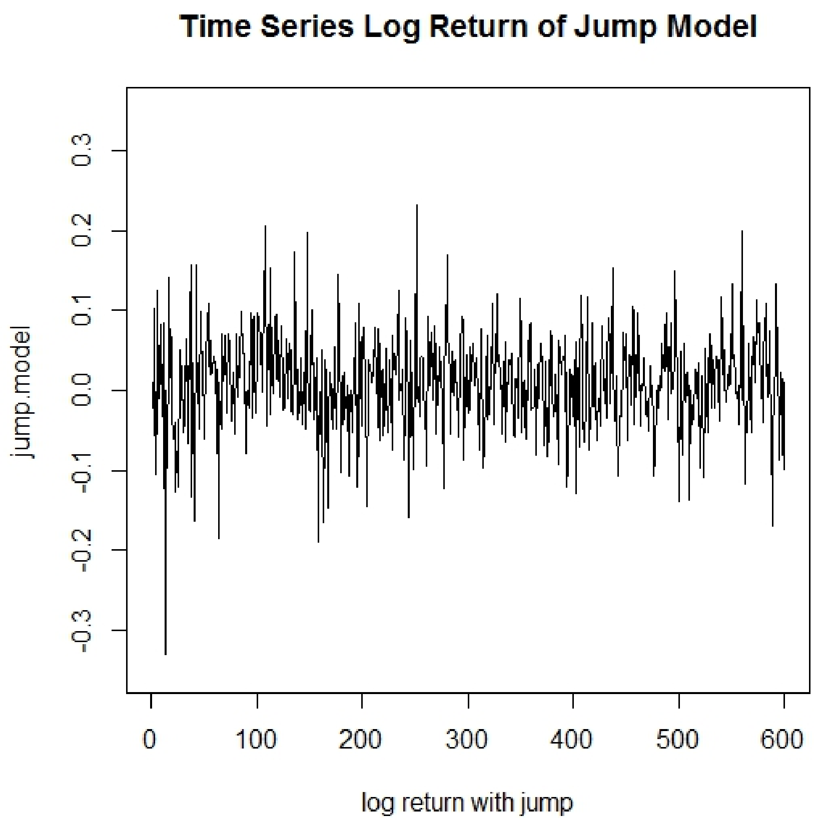

Figure 2: Time Series Log Return of the Jump Model

Problem 2: The Currency Carry Trade

The Deutsche Bank G10 Currency Future Harvest Index is, at any one time, composed of long futures contracts on the three G10 currencies associated with the highest interest rates and short futures contracts on the three G10 Currencies associated with the lowest interest rates.* The Index re-evaluates interest rates quarterly and, based on the evaluation, re-weights the futures contracts it holds. Immediately after each re-weighting, the Index will reflect an investment on a 2:1 leveraged basis in the three long and three short futures contracts (unless USD is one of the six currencies associated with the highest or lowest interest rates, in which case the Index will reflect an investment on a 1.66:1 leveraged basis). The PowerShares DB G10 Currency Harvest Fund tracks this index. Its ticker symbol is *DBV*. The price data is provided to you in the spreadsheet. Also provided in a second spreadsheet is the S&P500 index, whose ticker symbol is *GSPC*.

Table 1: Sample Moments

Mean	Std.Dev	Skewness	Kurtosis
0.007	0.065	-0.695	5.818

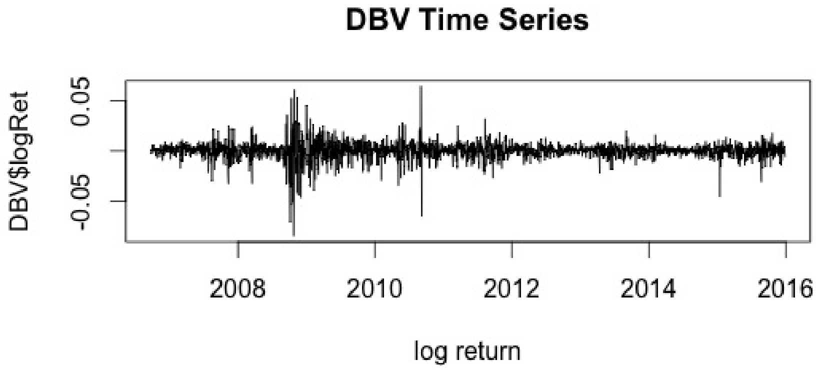

Figure 3: Time Series Log Return for DBV

Throughout this problem, use the adjusted closing values of the index. These values have adjusted for dividends and splits.

1. Visualizing the data.

   (a) Create time series plots of the daily log-returns for *DBV* and *GSPC*.

   Suggested solution: Shown in Graph 3 and Graph 5.

   (b) Create histograms of the daily log-returns for *DBV* and *GSPC*.

   Suggested solution: Shown in Graph 4 and Graph 6.

2. Shape of Return Distribution:

   (a) Test the null that the skewness of daily log returns is zero at the 5% significance level.

   Suggest Solution: T-statistic for skewness is $t = \frac{\hat{S}}{\sqrt{6/T}}$ where $\hat{S} = \frac{\frac{1}{n}\sum(X_i - \bar{X})^3}{[\frac{1}{n-1}\sum(X_i-\bar{X})^2]^{3/2}}$ is the sample skewness. We reject the null of $S = 0$ at a 5% level when $t > 1.96$. Results are shown in table 2.

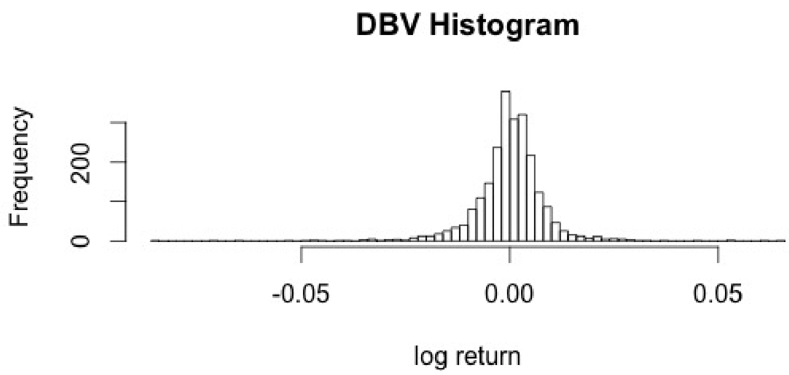

Figure 4: Histogram of the Log Return for DBV

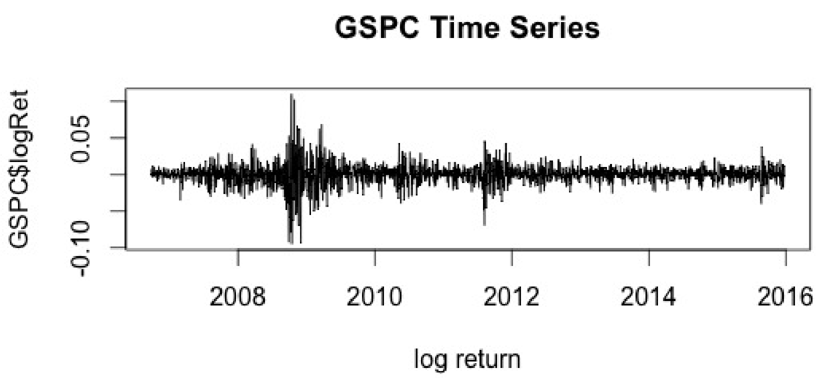

Figure 5: Time Series Log Return for GSPC

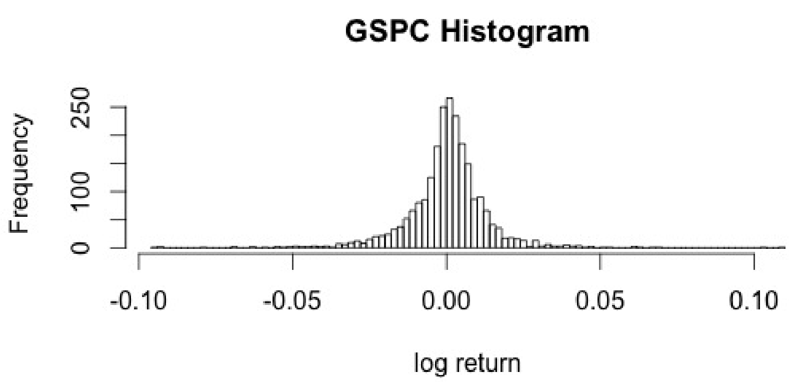

Figure 6: Histogram of the Log Return for GSPC

Table 2: Normality T-tests

DBV			GSPC		
Skewness	Kurtosis	JB	Skewness	Kurtosis	JB
-16.69	133.14	18004.85	-6.38	96.50	9354.26

(b) Test the null that the excess kurtosis of daily log returns is zero at the 5% significance level.

Suggest Solution: T-statistic for kurtosis is $t = \frac{\hat{K}-3}{\sqrt{24/T}}$ where $\hat{K} = \frac{\frac{1}{n}\sum(X_i-\bar{X})^4}{[\frac{1}{n-1}\sum(X_i-\bar{X})^2]^2}$ is the sample kurtosis. We reject the null of $K = 3$ at a 5% level when $t > 1.96$. Results are shown in table 2.

(c) Test the null that the daily log returns are normally distributed at the 5% significance level using the Jarque-Bera test.

Suggest Solution: JB-statistic is $t = \frac{(\hat{K}-3)^2}{24/T} + \frac{(\hat{S})^2}{6/T}$ where $\hat{K}$ and $\hat{S}$ are sample moments. We reject the null at a 5% level when $JB > 5.99$. Results are shown in table 2.

3. Compare all of these numbers in (a) and (b) to the same numbers for daily log returns on the S&P 500 measured over the same sample in one single table.

   Suggested solution: Table 2 shows the t-test results for skewness, kurtosis and JB normality test. We can easily reject the null at 5% level. Relatively speaking, DBV deviates from normal more.

4. Suppose you are a fund manager with a target return in mind (say 20 % per annum), and suppose that the ratio of expected returns to standard deviation is 0.50 for both investments. Using these numbers, discuss the different nature of the risks you would face if you invested in equity or currency markets to achieve that target return (with the appropriate leverage).

   Suggested solution: After comparing the index of the currency market and equity market, we can observe that for the currency market, the mean and variance are relatively low while the skewness and kurtosis are relatively high. This means that

Table 3: OLS and White OLS

	Estimates	Std.Err (OLS)	Std.Err (White)
β	0.431	0.010	0.018
Constant	-9.579e-05	1.361e-04	1.367e-04

if we are investing targeting same return and Sharpe ratio only, the currency market portfolio will be exposed to more higher moment risks.

5. Regress the log DBV returns on the log $GSPC$ returns. Report the standard errors of the slope and intercept coefficients using both standard OLS assumptions and allowing for non-normalities and heteroskedastic errors (White standard errors). Explain why the two are different and give the intuition for why White standard errors are smaller/larger than the classic OLS standard errors in this case.

   Suggested Solution: Table 3 shows the regression results. Note that the constant is not statistically significant. The White standard error is larger than the OLS standard error. This is because the squared value of x is positively correlated with the squared value of epsilon. Or, in other words, the conditional variance of the independent variable is positively correlated with the conditional variance of the residual. See "Postscript_PS1" on CCLE for more on this. This is intuitive because whenever the return of SP500 deviates too much, it means some extraordinary events happen to the equity market, which may not necessarily influence the currency market. Thus in these cases, the regression residuals will show larger variance.

Suggested R Codes for Problem 1

```
log.norm=rnorm(600,0.008,0.063)
hist(log.norm,breaks=100,main=Histogram of Log Normal Return")

J.norm=rnorm(600,-0.03,0.1)
r.norm=rnorm(600,0.012,0.05)
B=rbinom(600,1,0.15)

jump.model=r.norm+J.norm*B
hist(jump.model,breaks=100,main=Histogram of Jump Model")

sample.mean=mean(jump.model)
sample.variance=1/599*sum((jump.model-sample.mean)^2)
sample.skewness=1/600*sum((jump.model-sample.mean)^3)/sample.variance^(3/2)
sample.kurtosis=1/600*sum((jump.model-sample.mean)^4)/sample.variance^(2)
```

Suggested R Codes for Problem 2

```
DBV <- read.csv("~/Dropbox (Personal)/237E/Pset/DBV.csv")
DBV$lag=c(NA, DBV$Adj.Close[-nrow(DBV)])
DBV$ret=DBV$Adj.Close/DBV$lag
DBV$logRet=log(DBV$ret)
DBV$Date=as.Date(DBV$Date,"%m/%d/%y")
plot(DBV$Date,DBV$logRet,type="l")
hist(DBV$logRet,breaks=100)
DBV.n=nrow(DBV)
DBV.mean=mean(DBV$logRet[2:DBV.n])
DBV.var=1/(DBV.n-1)*sum((DBV$logRet[2:DBV.n]-DBV.mean)^2)
DBV.skew=1/DBV.n*sum((DBV$logRet[2:DBV.n]-DBV.mean)^3)/DBV.var^(3/2)
DBV.kurt=1/DBV.n*sum((DBV$logRet[2:DBV.n]-DBV.mean)^4)/DBV.var^(2)
DBV.t.skew=DBV.skew/((6/DBV.n)^0.5)
DBV.t.kurt=(DBV.kurt-3)/((24/DBV.n)^0.5)
DBV.JB=DBV.skew^2/(6/DBV.n)+(DBV.kurt-3)^2/(24/DBV.n)
```

```
#do the same thing for GSPC

reg=data.table(DBV$logRet[2:DBV.n],GSPC$logRet[2:GSPC.n])
library(DataAnalytics)
a=lm(V1~V2,data=reg)
library(sandwich)
vcovHC(a, type="HC0")
#gives the variance covarince matrix for the white-OLS
```