ECON 4140 N
Assignment 1
Ethan Wu 215551955
Jennifer Jennifer 216431934
Regina Mui 215674724

1. Define the random walk process.

The random walk process consists of Z_0, an arbitrary starting point, also the constant term; and Z_1, Z_2,...Z_t as IID (independent and identically distributed), which is the variable. The process shows the current value of a variable, including a past value from the previous time periods. It can be represented by

- Time 0: $S_0 = Z_0$
- Time 1: $S_1 = Z_0 + Z_1$
- Time t: $S_t = Z_0 + Z_1 + \ldots + Z_t$, where $t \geq 1$

$\therefore$ $S_0, S_1, \ldots, S_t$ is a random walk

According to the tutorial video, a random walk process defined by the formula:
Logpt = log(pt-1) +rt, where t is the time.

2. Does the time series of daily IBM prices display one trend or multiple trends?

Time series of daily ibm prices displays multiple trends

3. Given your answer to the first question, do prices behave like a stationary process or rather like a random walk?

Prices behave similar to random walk

4. Does the return series, defined as return(t) = log p(t)–log p(t-1) have a trend?

The return series defined as return(t) = log p(t) - log p(t-1) does not have a trend.

5. What are the differences between the behavior of prices and returns?

The behavior of prices are non-stationary and are similar to the random walk process, while the behavior of returns displays no trends and is stationary. This can be seen from observing the behavior of the return and price datas.

6. Define “volatility clustering”.

Volatility clustering is where large changes in price over time tend to occur and cluster together while when prices are calm, it tends to stay calm and have smaller changes.

7. Comment on the patterns in return and prices around July 2009 – what do you observe?

From observation 378 (July 1st) we could see that the closing price for IBM was $103.89 with a constant upward trend ending the month with observation 403 (July 31st) with the closing price of $116.86. We can also see a few instances of consolidation such as observation 381, 384 and 392, which promotes sustainability of growth for the ticker. The returns during this time horizon are also pretty consistent floating around 0.042 to - 0.030.

378	103.89	0.00405	378
379	100.81	-0.03010	379
380	100.73	-0.00079	380
381	99.28	-0.01450	381
382	99.76	0.00482	382
383	101.15	0.01384	383
384	99.91	-0.01233	384
385	102.68	0.02735	385
386	102.31	-0.00361	386
387	106.25	0.03779	387
388	109.63	0.03132	388
389	114.37	0.04233	389
390	115.38	0.00879	390
391	115.98	0.00519	391
392	114.52	-0.01267	392
393	116.00	0.01284	393
394	116.57	0.00490	394
395	116.56	-0.00009	395
396	116.21	-0.00301	396
397	116.19	-0.00017	397
398	116.79	0.00515	398
399	116.86	0.00060	399
400	118.83	0.01672	400
401	118.51	-0.00270	401
402	117.39	-0.00950	402
403	116.86	-0.00453	403

8. Are the IBM returns normally distributed? Explain and describe all evidence from the output provided by the summary statistics and figures (histogram, qq-plot, quantiles)

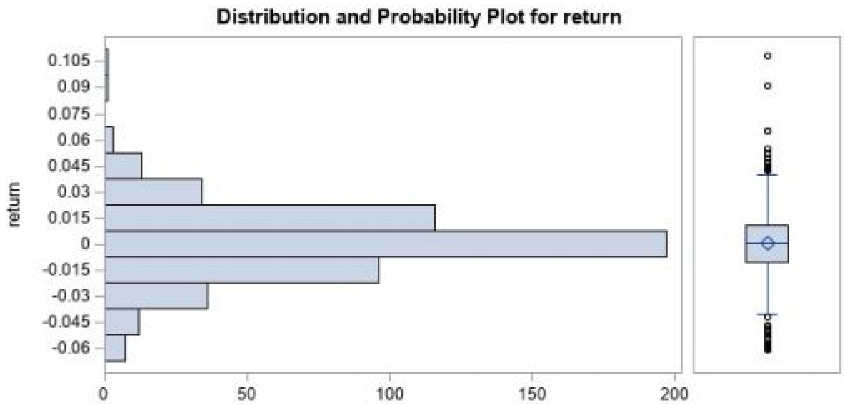
Skewness of 0.26566051 being very similar to 0 and kurtosis of 3.045312 being very close to 3 suggest that IBM returns are normally distributed

daily IBM prices 1/2/2008 - 1/20/2010

The UNIVARIATE Procedure
Variable: return

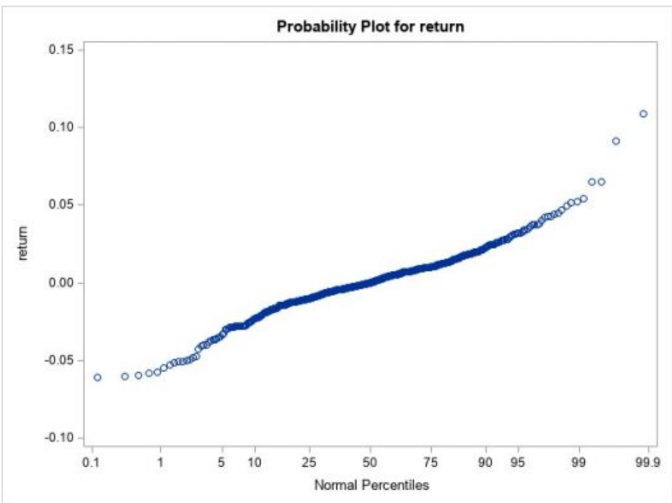
Moments			
N	516	Sum Weights	516
Mean	0.00049539	Sum Observations	0.25562312
Std Deviation	0.0200803	Variance	0.00040322
Skewness	0.26566051	Kurtosis	3.045312
Uncorrected SS	0.20778406	Corrected SS	0.20765743
Coeff Variation	4053.40213	Std Error Mean	0.00088399

The histogram and box and whisker plot below display almost perfect symmetry, which further supports IBM returns being normally distributed

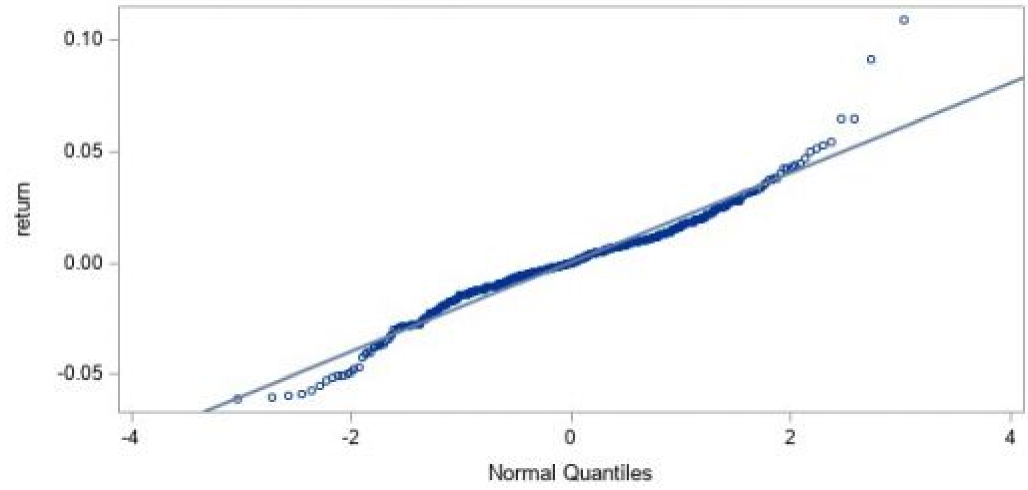


The P-P plot below forms an almost perfect straight line which also further supports normal distribution

The UNIVARIATE Procedure



Lastly, the Q-Q plot below almost perfectly aligns with the standard line further suggesting that the IBM return distribution is normal



9. Write the formula of the AR(1) process.

$Y_t = \varphi\,(Y_{t-1}\,) + \varepsilon_t$ when $|\,\varphi\,| < 1$

10. Explain the computer output and describe all evidence along the following lines:

10.1 what are the marginal mean and variance of returns? Is the mean of returns statistically significant – write the test statistic of the test for 0 mean and explain

From AR(1) process; $Y_t = \varphi\,(Y_{t-1}\,) + \varepsilon_t$
Marginal mean is $E(Y_t\,) = \boldsymbol{\mu}$ = Const / (1 - φ)

Variance of returns is 0.0004 , $1/\,(T\ -1)\ *\Sigma\,(rt\ -\ r\ bar)^2$
Mean is 0.0004, with t = 516

daily IBM prices 1/2/2008 - 1/20/2010

The UNIVARIATE Procedure
Variable: return

Moments			
N	516	Sum Weights	516
Mean	0.00049539	Sum Observations	0.25562312
Std Deviation	0.0200803	Variance	0.00040322
Skewness	0.26566051	Kurtosis	3.045312
Uncorrected SS	0.20778406	Corrected SS	0.20765743
Coeff Variation	4053.40213	Std Error Mean	0.00088399

Test for mean = 0;

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	0.560409	Pr > |t|	0.5754
Sign	M	0.5	Pr >= |M|	1.0000
Signed Rank	S	2615.5	Pr >= |S|	0.4394

- $H_0 : \mu = 0$
- $Ha : \mu \neq 0$

t = r bar / (√(σ head^2 / 516)
t = 0.56
Since Pr > |t| = 0.57 > 0.56,
We do not reject the null hypothesis H_0
Under the H_0:
t ≃ t (T-1)
α = 0.05 = (5%)
P-value = 0.5754

0.5754 > 0.05, thus we do not reject the H_0, that is $\mu = 0$

10.2 Are the returns serially correlated? Write the test statistic of the test of significance of the autoregressive coefficient and interpret. What is the estimated variance of the error term of the AR(1) model.

Test statistics of the test of significance of the autoregressive coefficient.

- If φ = 0:
  - H_0: φ = 0
  - H_1: φ ≠ 0
- If H_0 is rejected, returns are serially correlated and can be predicted
- If H_1 is not rejected, returns are serially uncorrelated and there is white noise

Estimates of Autoregressive Parameters			
Lag	Coefficient	Standard Error	t Value
1	0.058237	0.044033	1.32

Coefficient = ^ φ = 0.058237
t = ^ φ / S.D. (^ φ) = 0.058237/ 0.044033 = 1.32

If |t| > 1.96, we reject H_0,
however, t-value, 1.32 < 1.96

We accept H_0 where $\varphi = 0$
The error term in AR (1),ε : $r_t = \varepsilon_t$, $\varepsilon_t \sim W.N\ (0, \sigma_\varepsilon^2)$
∴ The returns are not serially correlated and predictable

10.3 What is the efficient market hypothesis?

The efficient market hypothesis (EMH) shows that random walk happens due to market efficiency. A market will be considered as 'efficient' if prices entirely reflect the available market information. The prices would remain the same when information is revealed to all participants of the market, and it is impossible to take advantage or make profits based on the information.

Information and price changes are unpredictable, which is why price changes will be random. Under the efficient market circumstance, only new information will affect the prices to change. In other words, there will be no reaction from the market when expected events occur.

There are three types of efficiency, defined with respect to the available information:

- Weak-form efficiency: information set of prices and returns history
- Semi-strong efficiency: information set that consists all publicly available information
- Strong-form efficiency: information set with information that all market participants know

10.4 Is your AR(1) estimation result consistent with the efficient market hypothesis? Explain why yes or not.

Yes, it is consistent with the efficient market hypothesis. The best prediction of future price is given pt and rt: that is the conditional mean (expectation)

$$\begin{aligned} E(\log p_t \mid \log p_t, r_t) &= \log p_t + E(r_{t+1} \mid r_t) \\ &= \log p_t + 0 \\ &= \log p_t \end{aligned}$$

Since the result from AR(1) is not serially correlated and predicted, thus the investor cannot predict the future price of IBM.

SAS code:

```
options linesize=78;

* read in the data *;

data ibm;
infile '\\tsclient\Drives\jenniferhanafi\Desktop\SAS\IBM.txt';
input price;
return = dif(log(price));
t=_n_;

* print the data *;

proc print data=ibm;
run;

* calculate the statistics *;

title 'daily IBM prices 1/2/2008 - 1/2/2010';

proc univariate normal plot;
var return;
probplot;

run;

* plot the prices *;

proc gplot;

plot price*t;

symbol interpol=joint;

run;


* plot the returns *;

title 'daily IBM returns 1/2/2008 - 1/2/2010';

proc gplot;

plot return*t;

symbol interpol=joint;
```

```
run;

* estimate the AR(1) model *;

proc autoreg;

model return =/nlag = 1;

run;
```