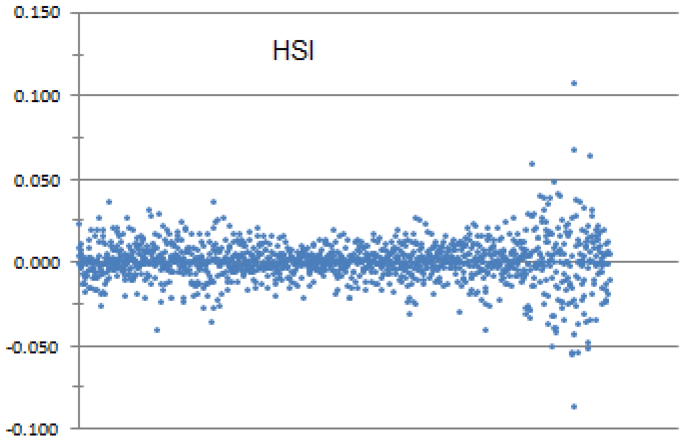
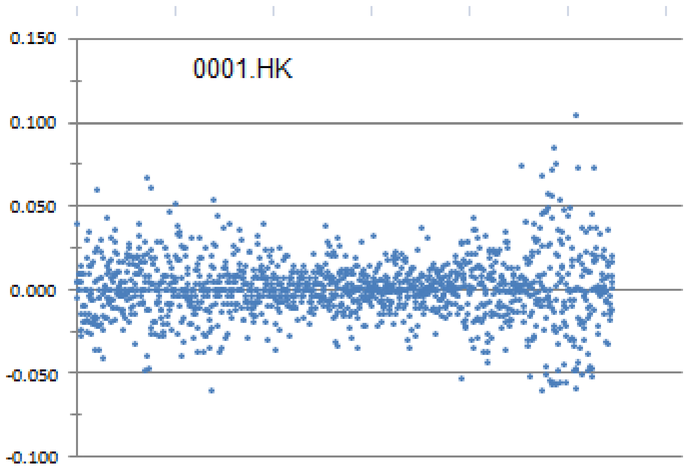
Validity of I.I.D assumption

I.I.D. random normal return, $r = \varepsilon(\mu, \sigma)$

(1) Is the probability distribution of asset price returns identical over time ? :





We see volatility clusterings.

i.e. periods of large returns are clustered and distinct from periods of small returns.

It is fair to think that volatility is changing with time (thus not identical).

(2) Are asset price returns statistically independent over time ? :

Check autocorrelation of historical data.

Given historical price returns $\{ r_1 , r_2 , \ldots , r_n \}$ with mean $E(r_t)$ and standard deviation $S(r_t)$.

Autocorrelation is the correlation with previous points.

Consider the covariance between $\{ r_1 , r_2 , \ldots , r_{n-k} \}$ and $\{ r_{1+k} , r_{2+k} , \ldots , r_n \}$ for fixed k.

i.e. covariance of pairs $(r_1 , r_{1+k}) , (r_2 , r_{2+k}) , \ldots , (r_{n-k} , r_n)$

Autocorrelation coefficient of time lag k, $\rho_k = \dfrac{E\Big(\big(r_t - E(r_t)\big)\big(r_{t+k} - E(r_t)\big)\Big)}{S^2(r_t)}$

If $\{ r_1 , r_2 , \ldots , r_n \}$ is not autocorrelated then ρ_k will not be significantly different from zero.

Refer to autocorrelation_daily.xls

k	^HSI	0001.HK	0002.HK	0003.HK	0004.HK	0005.HK
1	-0.03332	0.02798	-0.14105	-0.09228	0.02387	-0.08025
2	-0.03019	-0.03035	-0.00790	0.00168	-0.06643	0.01870
3	-0.00920	-0.03341	-0.00270	-0.01095	-0.02018	-0.08118
4	0.00142	-0.05717	0.02490	-0.02416	-0.00471	0.05124
5	-0.04178	-0.04280	-0.02634	-0.05486	-0.04795	-0.01924
6	0.01358	0.00614	-0.06278	-0.04324	-0.01008	0.01575
7	0.01765	-0.01358	0.01502	-0.01512	-0.02719	0.01187
8	-0.01116	-0.01738	-0.06277	-0.00865	-0.03733	0.00501
9	-0.02365	-0.00773	0.02110	0.04165	0.03850	-0.01730
10	-0.02027	0.02996	-0.02713	-0.06544	-0.00968	-0.04735

Weak autocorrelation in daily returns.

Random (*i.e.* independent) normal return is still valid, $r_t = \varepsilon(\mu , \sigma_t)$

Note : In general, μ is small compared with σ. For simplicity, we ignore the dynamical behaviour of μ and consider it to be a constant.

Autocorrelation of squared returns :

Consider the covariance between { r_1^2 , r_2^2 , … , r_{n-k}^2 } and { r_{1+k}^2 , r_{2+k}^2 , … , r_n^2 } for fixed k.

i.e. covariance of pairs (r_1^2 , r_{1+k}^2) , (r_2^2 , r_{2+k}^2) , … , (r_{n-k}^2 , r_n^2)

Autocorrelation coefficient of time lag k, $\rho_k = \dfrac{E\left(\left(r_t^2 - E(r_t^2)\right)\left(r_{t+k}^2 - E(r_t^2)\right)\right)}{S^2(r_t)}$

Refer to autocorrelation_daily.xls

Autocorrelation ρ_k :

k	^HSI	0001.HK	0002.HK	0003.HK	0004.HK	0005.HK
1	0.372590	0.260961	0.412046	0.134193	0.225554	0.411874
2	0.326622	0.293218	0.083010	0.163164	0.154001	0.187832
3	0.281797	0.191595	0.112685	0.120197	0.095349	0.194677
4	0.242100	0.168220	0.139903	0.117063	0.080983	0.196324
5	0.206796	0.172383	0.124422	0.070604	0.093043	0.142218
6	0.137057	0.165413	0.096127	0.099879	0.095991	0.056128
7	0.150806	0.158780	0.135060	0.107453	0.051393	0.085659
8	0.171524	0.202179	0.119837	0.061247	0.050559	0.106649
9	0.116047	0.137808	0.107984	0.122715	0.032207	0.052696
10	0.186361	0.140596	0.147379	0.102280	0.028836	0.115942

Strong autocorrelation in squared daily return.

In the model, $r_t = \varepsilon(\ \mu\ ,\ \sigma_t\ )$, or $(r_t - \mu) = \sigma_t\ \varepsilon$

Thus, $(r_t - \mu)^2 = \sigma_t^2 \varepsilon^2$ and $(r_{t+k} - \mu)^2 = \sigma_{t+k}^2 \varepsilon^2$

It should be noted that the two $\boldsymbol{\varepsilon}$ above are independent.

If there are strong autocorrelation between r_t^2 and r_{t+k}^2 , this implies a strong autocorrelation between σ_t^2 and σ_{t+k}^2 .

1. Given price returns $\{ r_1 , r_2 , \dots , r_t \}$,

We consider the future return $r_{t+1} = \varepsilon(\mu , \sigma_{t+1}) = \mu + \sigma_{t+1}\,\varepsilon$ with constant mean μ.

$\boldsymbol{\varepsilon}$ is an independent random normal number. σ_{t+1} is also considered to be a random quantity independent of the current $\boldsymbol{\varepsilon}$ but it is correlated with previous ($\sigma_t , \sigma_{t-1} , \dots$).

$E(r_{t+1}) = \mu + E(\sigma_{t+1}\,\varepsilon)$

$= \mu + E(\sigma_{t+1})\, E(\varepsilon)$ by the independence between σ_{t+1} and ε

$= \mu$ (This means that if we can repeat the generation of r_{t+1} many times, we can estimate μ very accurately. However, this may not be possible in reality)

We also have $E(r_1) = E(r_2) = \dots = E(r_t) = \mu$

Thus, $r_1 + r_2 + \dots + r_t = t\mu + \sigma_1\,\varepsilon_1 + \sigma_2\,\varepsilon_2 + \dots + \sigma_t\,\varepsilon_t$

$E(r_1 + r_2 + \dots + r_t) = t\mu + E(\sigma_1\,\varepsilon_1) + E(\sigma_2\,\varepsilon_2) + \dots + E(\sigma_t\,\varepsilon_t) = t\mu$

This gives $\mu = (1/t)\, E(r_1 + r_2 + \dots + r_t) \cong (1/t)\,(r_1 + r_2 + \dots + r_t)$, unbiased estimation

(This means that we can extract one set of $\{ r_1 , r_2 , \dots , r_t \}$ in order to estimate μ accurately.)

2. Consider the model $\sigma_{t+1}^2 = \omega + \alpha(r_t - \mu)^2 + \beta\sigma_t^2$ with weight factors $\alpha > 0$, $\beta > 0$, and $\omega > 0$.

$E(\sigma_{t+1}^2) = \omega + \alpha\, E((r_t - \mu)^2) + \beta\, E(\sigma_t^2)$

Since $(r_t - \mu)^2 = \sigma_t^2\varepsilon_t^2 \;\Rightarrow\; E((r_t - \mu)^2) = E(\sigma_t^2\varepsilon_t^2) = E(\sigma_t^2)\, E(\varepsilon_t^2) = E(\sigma_t^2)$

Thus, $E(\sigma_{t+1}^2) = \omega + (\alpha + \beta)\, E(\sigma_t^2)$

Stationary condition requires all the expected values to be equal, $E(\sigma_1^2) = E(\sigma_2^2) = \dots = V_L$

This gives $E(\sigma_t^2) = \omega/(1 - \alpha - \beta) \;\Rightarrow\; (\alpha + \beta) < 1$

also, $\omega = (1 - \alpha - \beta)\, E(\sigma_t^2) = (1 - \alpha - \beta)\, V_L$

It is also true that $E((r_1 - \mu)^2) = E((r_2 - \mu)^2) = \dots = V_L$

So, $E((r_1 - \mu)^2 + (r_2 - \mu)^2 + \dots + (r_t - \mu)^2) = E((r_1 - \mu)^2) + E((r_2 - \mu)^2) + \dots + E((r_t - \mu)^2)$

$= t\, V_L$

or, $V_L = (1/t)\, E((r_1 - \mu)^2 + (r_2 - \mu)^2 + \dots + (r_t - \mu)^2)$

$\cong (1/t)[\, (r_1 - \mu)^2 + (r_2 - \mu)^2 + \dots + (r_t - \mu)^2 \,]$, unbiased estimation