Understanding variations in financial returns

Susan Thomas

10 March, 2011

Goals

- Returns variations
- Dynamics of risk

Part I

Features of variation in returns

Returns

- Returns are calculated as log price differences and expressed in %:

$$r = 100 * diff(log(P_t))$$

- Mapping from prices to returns (or vice-versa) should be trivial.

$$\begin{aligned} r_1 &= 100 * log(P_1/P_0) \\ P_1 &= P_0 e^{r_1/100} \end{aligned}$$

- Standard assumption: returns come from a known unconditional distribution, $f(r)$.
-

Returns

- Returns are calculated as log price differences and expressed in %:

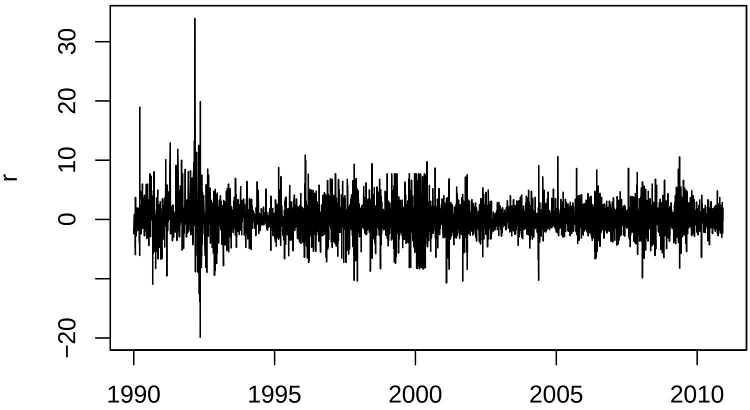
$$r = 100 * diff(log(P_t))$$

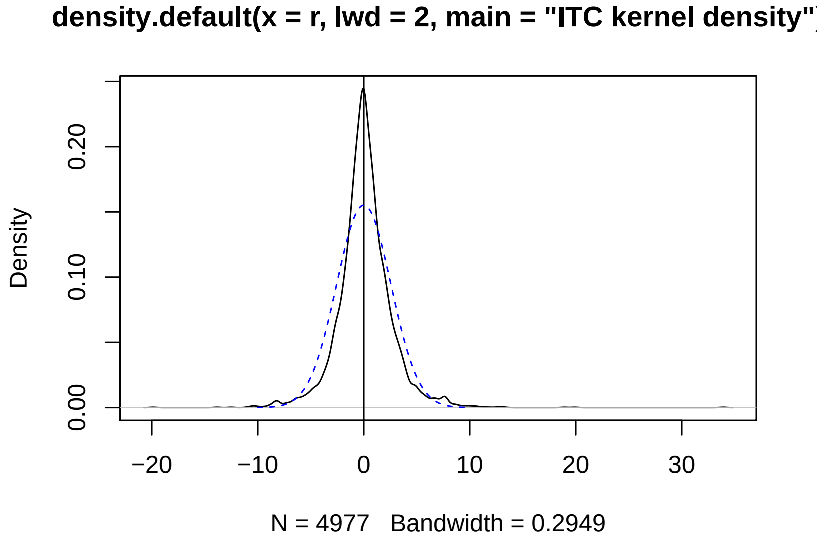
- Mapping from prices to returns (or vice-versa) should be trivial.

$$\begin{aligned} r_1 &= 100 * log(P_1/P_0) \\ P_1 &= P_0 e^{r_1/100} \end{aligned}$$

- Standard assumption: returns come from a known unconditional distribution, $f(r)$.
-

returns




density.default(x = r, lwd = 2, main = "ITC kernel density")
Density
0.00
0.10
0.20
−20
−10
0
10
20
30
N = 4977 Bandwidth = 0.2949

Non-normal distribution

- We see that in the tails, the normal distribution lies below the data histogram.
- There is (slight) asymmetry in the right and left side of the data histogram.
- Reasons?
  1. Real distribution is non-normal with fatter tails?
  2. Data comes from a mix of distributions: normal with low risk + normal with high risk?
  3. Real distribution is non-stationary?

Non-normal distribution

- We see that in the tails, the normal distribution lies below the data histogram.
- There is (slight) asymmetry in the right and left side of the data histogram.
- Reasons?
  1. Real distribution is non-normal with fatter tails?
  2. Data comes from a mix of distributions: normal with low risk + normal with high risk?
  3. Real distribution is non-stationary?

Different distribution?

- One candidate: Student's t-distribution with low degrees of freedom.
- The possibility of seeing large deviations are:

	Tail prob.		Fraction days in a year	
Deviation	Normal	t(df = 4)	Normal	t(df = 4)
-5%	0.0000	0.0038	0.00	0.94
-4%	0.0000	0.0081	0.01	2.02
-3%	0.0014	0.0200	0.34	4.99
-2%	0.0228	0.0581	5.69	14.51
-1%	0.1587	0.1870	39.66	46.74

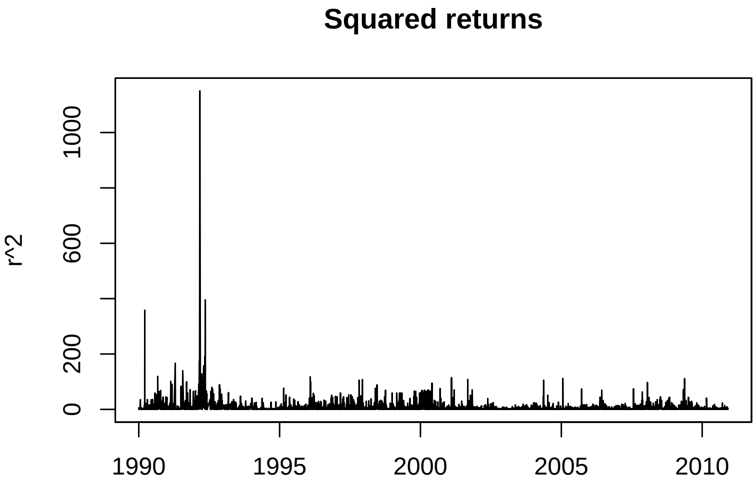
Alternative distribution comparisons

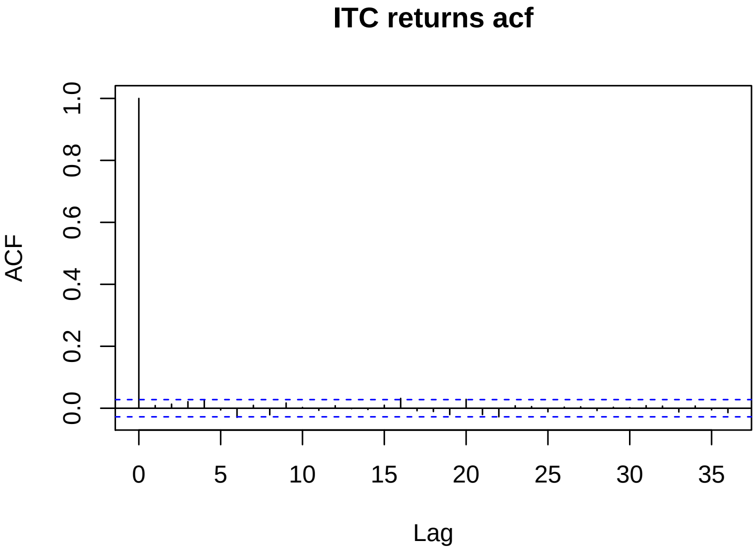
Our data: Sample size = 4977

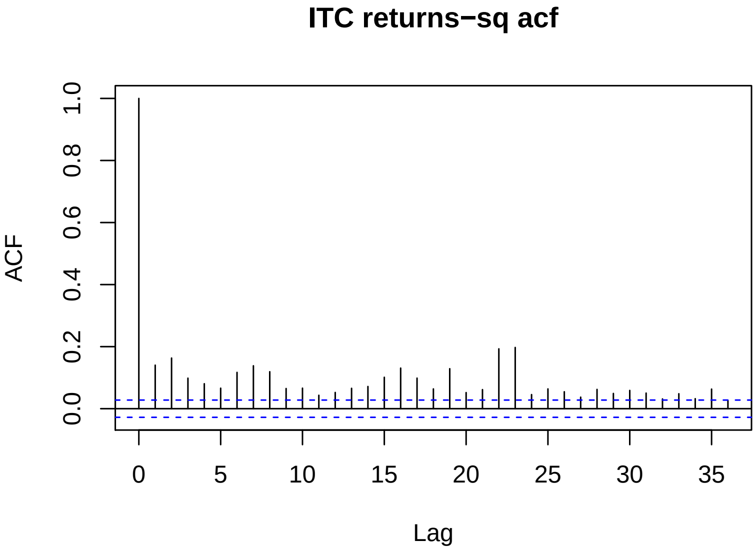
	-5	-4	-3	-2	-1
Normal	0.0000	0.0000	0.0014	0.0228	0.1587
t(df=4)	0.0038	0.0081	0.0200	0.0581	0.1871
Data	4e-04	0.0014	0.0078	0.0247	0.1039

Time varying variances

- An alternative explanation is that volatility changes every day: heteroskedasticity.
- This implies that the observed data comes out of a mixture of distributions.


Squared returns
r^2
0
200
600
1000
1990
1995
2000
2005
2010


ITC returns acf
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
35
Lag


ITC returns−sq acf
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
35
Lag

Summary stats for I.T.C. Ltd.

- Mean = 0.11%
- Std. Dev. = 2.57%
- AR model order on returns = 0
- AR model order on volatility (squared returns) = 25
- There is a strong case to believe that returns come from a mixture of distributions: normal with different σ.

Choices in modelling variations in σ

- Homoskedasticity: Constant unconditional variance
- Key idea: conditional heteroskedasticity that is autoregressive

$$\begin{aligned} r_t &= \text{arma(p,q)} + \epsilon_t \\ \epsilon_t &\sim N(\mu, h_t) \\ h_t &= \gamma_0 + \gamma_1 \epsilon_{t-1}^2 + \ldots + \gamma_k \epsilon_{t-k}^2 \end{aligned}$$

- Joint model of returns and volatility dynamics; source of information is returns itself.
- Concise version: GARCH(k,m)

$$h_t = \gamma_0 + \gamma_1 \epsilon_{t-1}^2 + \gamma_k \epsilon_{t-k}^2 + \beta_1 h_{t-1} + \ldots + \beta_m h_{t-m}$$

- Simplest version: GARCH(1,1)

$$h_t = \gamma_0 + \gamma_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}$$

ARCH implications on VaR

- Unconditional variance:

$$\text{arch} : h = \gamma_0/(1 - \sum_{i=1}^{i=k} \gamma_i)$$

$$\text{garch} : h = \gamma_0/(1 - \sum_{i=1}^{i=k} \gamma_i - \sum_{j=1}^{j=m} \beta_i)$$

Example: GARCH(1,1) $h = \gamma_0/(1 - \gamma_1 - \beta_1)$

- *Inference*: estimated historical σ is an underestimate of actual σ.
- **Note**: Since $\sigma_1 = f(\sigma_0)$, variance of data aggregated over K periods no longer follows the $\sqrt{K}$ rule.
- **Note**: When $\gamma_1 + \beta_1 = 1$, $h \to \infty$ which is a non-stationary process.
- **Note**: $\gamma_1 + \beta_1$ measures the *persistence* of variance.

ARCH implications on VaR

- Unconditional variance:

$$\text{arch}: h = \gamma_0/(1 - \sum_{i=1}^{i=k} \gamma_i)$$

$$\text{garch}: h = \gamma_0/(1 - \sum_{i=1}^{i=k} \gamma_i - \sum_{j=1}^{j=m} \beta_i)$$

Example: GARCH(1,1) $h = \gamma_0/(1 - \gamma_1 - \beta_1)$

- *Inference*: estimated historical σ is an underestimate of actual σ.
- **Note**: Since $\sigma_1 = f(\sigma_0)$, variance of data aggregated over K periods no longer follows the $\sqrt{K}$ rule.
- **Note**: When $\gamma_1 + \beta_1 = 1$, $h \rightarrow \infty$ which is a non-stationary process.
- **Note**: $\gamma_1 + \beta_1$ measures the *persistence* of variance.

ARCH implications on VaR

- Unconditional variance:

$$\text{arch}: h = \gamma_0/(1 - \sum_{i=1}^{i=k} \gamma_i)$$

$$\text{garch}: h = \gamma_0/(1 - \sum_{i=1}^{i=k} \gamma_i - \sum_{j=1}^{j=m} \beta_i)$$

Example: GARCH(1,1) $h = \gamma_0/(1 - \gamma_1 - \beta_1)$

- *Inference*: estimated historical σ is an underestimate of actual σ.
- **Note**: Since $\sigma_1 = f(\sigma_0)$, variance of data aggregated over K periods no longer follows the $\sqrt{K}$ rule.
- **Note**: When $\gamma_1 + \beta_1 = 1$, $h \to \infty$ which is a non-stationary process.
- **Note**: $\gamma_1 + \beta_1$ measures the *persistence* of variance.

Comparing popularly used ARCH models

- EWMA: Equally Weighted Moving Average

$$
\begin{aligned}
h_t &= \lambda h_{t-1} + (1-\lambda)\epsilon^2_{t-1} \\
&= \lambda^2 h_{t-2} + (1-\lambda)^2\epsilon^2_{t-2} + +(1-\lambda)\epsilon^2_{t-1} \\
&= \lambda^3 h_{t-3} + (1-\lambda)^2\epsilon^3_{t-3} + (1-\lambda)^2\epsilon^2_{t-2} + (1-\lambda)\epsilon^2_{t-1} \\
&= \sum_{j=1}^{\infty}(1-\lambda)^j\epsilon^2_{t-j}
\end{aligned}
$$

- Most famous EWMA: *RiskMetrics*, $\lambda = 0.94$
- GARCH with $\gamma_0 = 0$ and $\gamma_1 + \beta_1 = 1$.
- On short time horizons, the forecasts of the two models tend to be the same.

Garch for I.T.C Ltd.

	Estimate	Std. Error	t value	$\Pr(>\|t\|)$
μ	0.1169	0.0293	3.992	6.55e-05***
ϕ_1	0.0252	0.0155	1.625	0.104
γ_0	0.1895	0.0309	6.131	8.74e-10***
γ_1	0.0976	0.0099	9.869	$< 2e-16^{***}$
β_1	0.8748	0.0127	69.178	$< 2e-16^{***}$

Testing