University of Akron
Department of Statistics

Topics in Statistical Inference Regarding GARCH Model

Nao Mimoto [1]

[1] nmimoto@uakron.edu

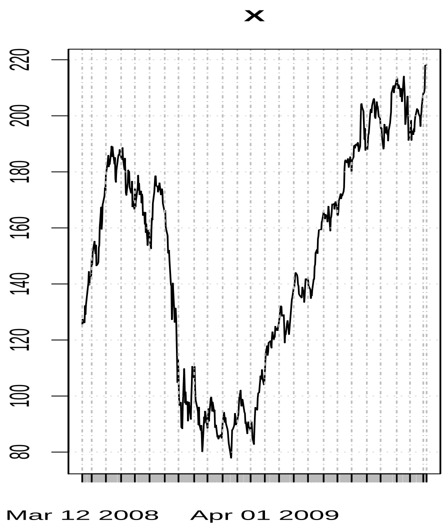
Contents

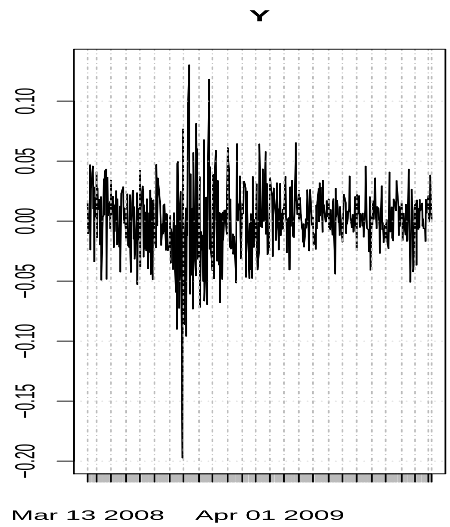
1

GARCH(p,q) Model

Daliy Return and Volatility

Aplle stock price 2008-03-12 to 2010-03-08

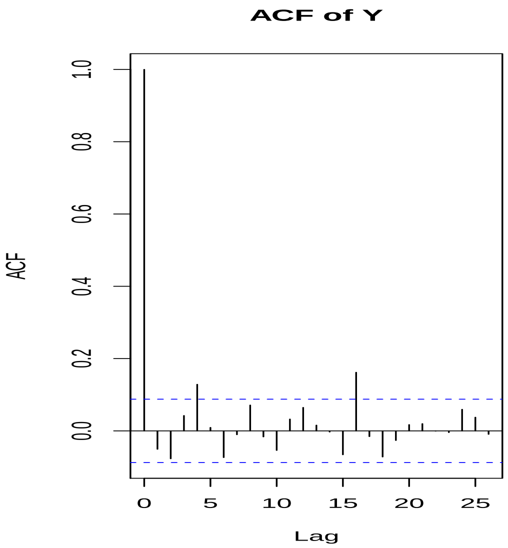


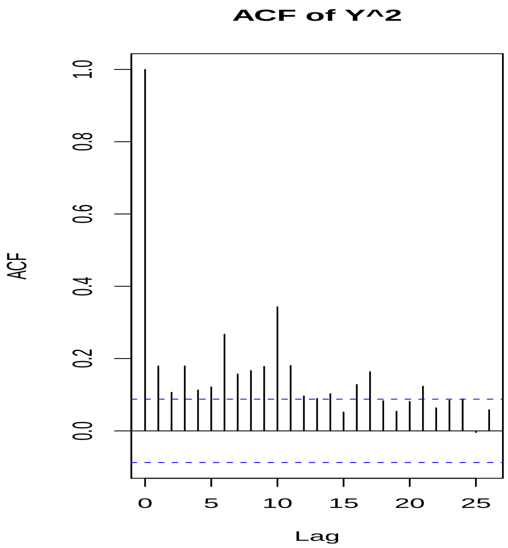


X_t

$Y_t = \log(X_t) - \log(X_{t-1})$

Autocorrelation of daily return





ACF of Y_t ACF of Y_t^2

```
---Rcode---
require('quantmod')
getSymbols("AAPL") #download from Yahoo! Finance

X <- AAPL$AAPL.Adjusted[300:800]; plot(X)  #2008-03-13 to 2010-03-08
Y <- diff(log(X))[-1]; plot(Y)

layout(matrix(c(1,2), 1, 2, byrow = T)); acf(as.ts(Y), main="ACF of Y"); acf(as.ts(Y^2), main="ACF of Y^2")

n="APPL-2"; dev.copy(postscript, paste(n,'.eps', sep=""), horizontal=F); dev.off();
dev.copy(pdf,paste(n,'.pdf',sep="")); dev.off();
```

The GARCH model

- Observations: $y_1, \ldots, y_n$

- GARCH(p,q) model

$$\begin{aligned} y_k &= \sigma_k \epsilon_k \\ \sigma_k^2 &= \omega + \sum_{i=1}^{p} \alpha_i y_{k-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{k-j}^2 \end{aligned}$$

- ϵ_i are i.i.d r.v. with mean 0 and variance 1 with density f.

- The parameter $\boldsymbol{\theta} = (\omega, \alpha_1, \beta_1)$, with $\omega > 0$, and $\alpha_1, \beta_1 \geq 0$.

- Assume θ satisfy the sufficient condition for the stationarity e.g. for GARCH(1,1)

$$\alpha_1 + \beta_1 < 1$$

Problem:

Want to perform goodness-of-fit test for the distribution of ϵ_i, possibly identify it.

- Errors are typically not Gaussian.
- for better parameter estimation.
- for better order selection.
- for better model selection.
- to idenfity skewed error distribution (and skewed marginal distribution).
- for more accurate Monte Carlo Simulation.
- for better forecasting.

Estimationg θ

- Quasi-Maximum Likelihood Estimator: Assume normality, and use MLE. The estimator is consistent and asymptotocally normal even under the non-normal error distribution.

- MLE is better whenever possible (c.f. Engle and González-Rivera (1991), González-Rivera and Drost (1999), Berkes and Horváth (2004), Francq and Zakoian (2004))

Q-MLE vs MLE

Estimating parameter $\boldsymbol{\theta} = \{1, .1, .8\}$ when f is T(5), n=500

SD of estimates	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
MLE	0.5728	0.0334	0.0812
QMLE	0.7642	0.0507	0.1042
Increase	25%	51%	28%

Order Selection

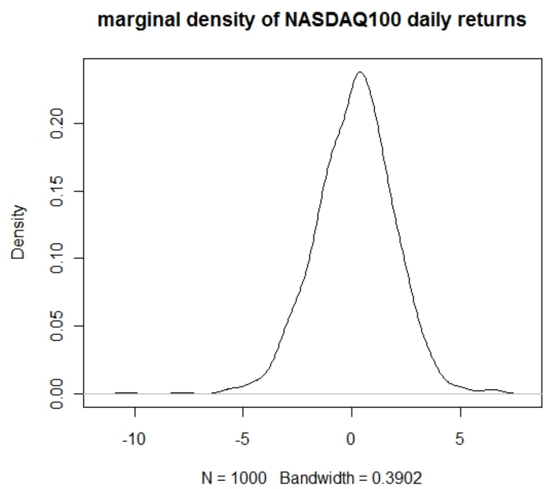
Use AIC to identify the GARCH(1,1) model from the choice of GARCH(p, q) models with $1 \leq p, q \leq 4$, (n=1000, $\boldsymbol{\theta} = \{1, .1, .8\}$, iteration=1000)

True f	likelihood used	% of Slelecting Correct Order
Normal	Normal	70.4%
T(5)	Norm	21.4%
T(5)	T(5)	72.5%

Conditional Skewness

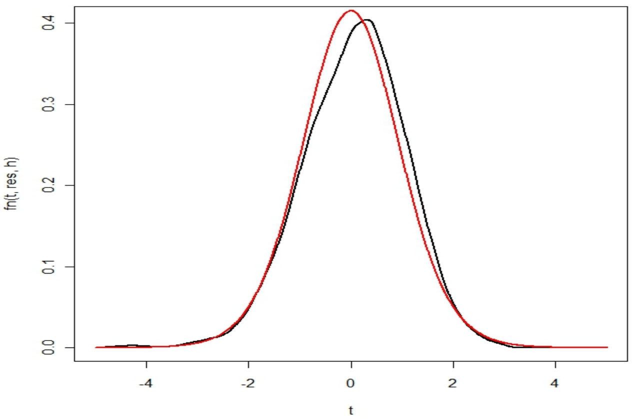
Daily log return of NASDAQ100 from 1995 Aug 18, to 1999 Aug 4. (1000 obs)

- mean 0.1321
- skewness -0.29



Conditional Skewness

- GARCH(1,1) fitted to the data.
- $\hat{\boldsymbol{\theta}} = (.277,\ .135,\ .782)$, residuals passes randomness test.
- residual density Compared to standardized T(10) density.
- Is the difference significant? Should skewed distribution be used for errors?



Theoretical Result

Goodness-of-Fit test under i.i.d.

- Kolmogorov-Smirnov test looks at maximum difference between CDF and Empirical CDF.

- KS test is non-parametric. i.e. Can be used to any specified null hypothesis.

- The test is distribution free. i.e. Distribution of the test statistic does not depend on the choice of the null hypothesis.

- When the parameters of the null hypothesis has to be estimated, the test is no longer distribution free, and the critical value must be calculated via simulation.

- Bickel and Rosenblatt (1974) proposed a test based on the Kernel Density Estimator that is asymptotically distribution free even when the parameters are estimated.

Kernel density estimator

- If i.i.d. errors $\epsilon_1, \ldots, \epsilon_n$ are observable,

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\Big(\frac{x - \epsilon_i}{h_n}\Big)$$

- Only residuals $\hat{\epsilon}_i$ are available.

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\Big(\frac{x - \hat{\epsilon}_i}{h_n}\Big)$$

Test with Kernel Density Estimator

Bickel and Rosenblatt (1974) and Bachmann and Dette (2002),

- As $n \to \infty$, let $h \to 0$ and $nh^2 \to \infty$.

$$n\sqrt{h} \int (f_n(x) - Ef_n(x))^2 dx - \frac{m}{\sqrt{h}} \quad \to_D \mathcal{N}(0, \tau^2)$$

  where $m = \int K^2(t)dt$ and
  $\tau^2 = 2 \int f^2(x)dx \int (K * K)^2(x)dx$.

- Does this result hold for GARCH model errors?

How residuals are calculated

- innovations:
$$\epsilon_k = y_k/\sigma_k,$$
- With pilot estimator $\boldsymbol{\theta}_n$:
the residuals:
$$\hat{\epsilon}_k = y_k/\hat{\sigma}_k.$$
- (cf. Berkes, Horváth and Kokoszka (2003)) Represent the conditional volatility using past observations
$$\sigma_k^2 = w_k(\boldsymbol{\theta}) = c_0(\boldsymbol{\theta}) + \sum_{i=1}^{\infty} c_i(\boldsymbol{\theta}) y_{k-i}^2$$

Representation of σ_k^2

- Representation the conditional volatility using past observations

$$\sigma_k^2 = w_k(\boldsymbol{\theta}) = c_0(\boldsymbol{\theta}) + \sum_{i=1}^{\infty} c_i(\boldsymbol{\theta}) y_{k-i}^2$$

- Estimate σ_k^2 by its truncated version with parameter estimator:

$$\hat{\sigma}_k^2 = \tilde{w}_k(\boldsymbol{\theta}_n) = c_0(\boldsymbol{\theta}_n) + \sum_{i=1}^{k-1} c_i(\boldsymbol{\theta}_n) y_{k-i}^2$$

- Need pilot estimator θ_n.
- Assume $\sqrt{n}$-consistency of $\boldsymbol{\theta}_n$

$$\sqrt{n}(\boldsymbol{\theta}_n - \boldsymbol{\theta}) = O_p(1)$$

- Example: Q-MLE

Assumptions

- K is bounded symmetric density on $[-1,1]$, twice differentiable with bounded derivative K' and $\int (K''(z))^2 dz < \infty$.

- Fisher information for location and scale parameters in one observation from f are finite. (implies that f is bounded and Lipschitz (1/2).)

- $\int \int x^2 [\dot{f}(x - zh) - \dot{f}(z)]^2 dx K(z) dz \to 0$ as $h \to 0$.

- Assumption includes double exponential.

- $E|\epsilon_0^2|^\delta < \infty$ for $\delta > 1$.

Goodness-of-fit test

$$H_0 : f = f_0 \quad vs \quad H_A : H_0 \text{ is false}$$

As $n \to \infty$, $h \to 0$ and $nh^5 \to \infty$.
Then

$$n\sqrt{h} \int (\hat{f}_n(x) - Ef_n(x))^2 dx - \frac{m}{\sqrt{h}} \quad \to_D \mathcal{N}(0, \tau^2)$$

where $m = \int K^2(t)dt$ and
$\tau^2 = 2 \int f^2(x)dx \int (K * K)^2(x)dx$.

- The test is asymptotically distribution-free.
- You can use this test to test for any specified f_0.

Simulation Study

Simulation parameters

For finite n (< 1000) we want to investigate the test statistic distribution's:

- dependence on $\boldsymbol{\theta}$
- dependence on n
- dependence on f (f_0)
- dependence on h

Simulation parameters

- $\boldsymbol{\theta} = (1, .1, .8)$

- For each iteration, $\boldsymbol{\theta}_n$ was obtained using Q-MLE.

- Iterated 2000 times.

- Epanechnikov kernel $K(u) = (3/4)(1 - u^2)I(|u| \leq 1)$

- Bandwidth
$$h = \left(\int K^2(x)dx \Big/ \int (f_0''(x))^2 dx (\int x^2 K(x)dx)^2\right)^{1/5} n^{-1/t}$$
with $t = 5.1$. If $t = 5$ is an optimum bandwidth which minimizes the AMISE under iid case.

Test for normality

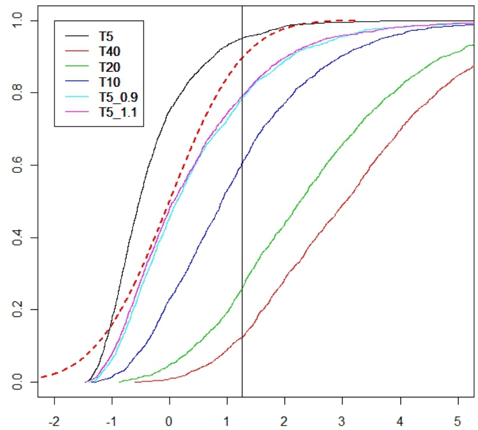
$$H_0 : f = \text{ is N(0,1)} \quad vs \quad H_A : H_0 \text{ is false}$$

n	$f\backslash$ Tests	$\widehat{T}_{emp}$	$\widetilde{T}_{emp}$	KS	JB	$\widehat{T}_{asy}$	$\widetilde{T}_{asy}$
	f_0 =N	0.050	0.050	0.050	0.050	0.003	0.016
	T40	0.058	0.032	0.060	0.126	0.006	0.010
	T20	0.094	0.025	0.060	0.335	0.007	0.008
500	T10	0.298	0.042	0.116	0.734	0.064	0.012
	T5	0.916	0.524	0.505	0.994	0.704	0.368
	$T5_{0.9}$	0.946	0.652	0.610	0.996	0.794	0.507
	$T5_{1.1}$	0.952	0.651	0.593	0.996	0.768	0.486
	Lap	1.000	0.996	0.984	1.000	0.998	0.987
	Logi	0.566	0.102	0.196	0.879	0.194	0.040

- Can't beat the Jaque-Bera test of normality.

$\widehat{T}$ test for T5

n=500
simulated critical value for size .05: 1.25

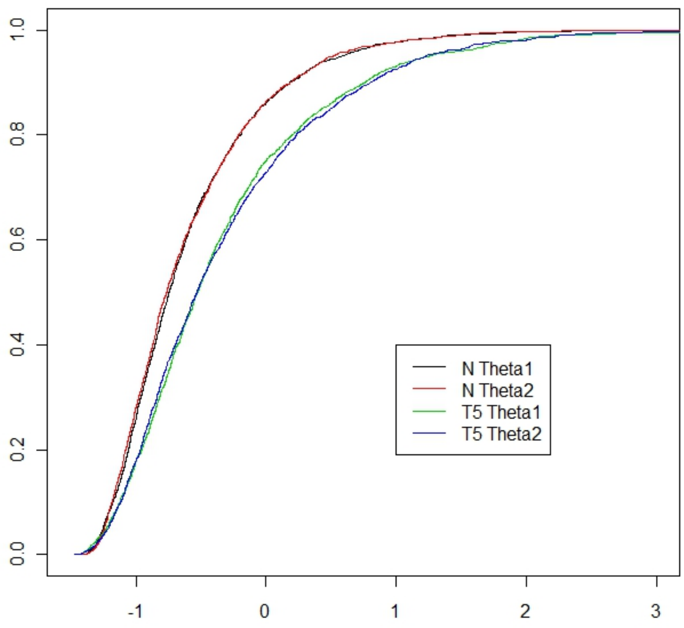


Test for T5

n	$f\backslash$ Tests	$\widehat{T}_{emp}$	$\widetilde{T}_{emp}$	KS	$\widehat{T}_{asy}$	$\widetilde{T}_{asy}$
500	f_0 =T5	0.050	0.050	0.050	0.018	0.062
	T40	0.876	0.949	0.434	0.725	0.964
	T20	0.744	0.868	0.330	0.572	0.888
	T10	0.396	0.568	0.182	0.236	0.609
	$T5_{0.9}$	0.216	0.170	0.146	0.117	0.198
	$T5_{1.1}$	0.210	0.164	0.147	0.106	0.188
1000	f_0 =T5	0.050	0.050	0.050	0.028	0.076
	T40	0.996	1.000	0.790	0.989	1.000
	T20	0.973	0.995	0.658	0.946	0.998
	T10	0.670	0.807	0.304	0.568	0.878
	$T5_{0.9}$	0.398	0.310	0.260	0.318	0.406
	$T5_{1.1}$	0.308	0.236	0.211	0.237	0.321

Dependence on $\boldsymbol{\theta}$

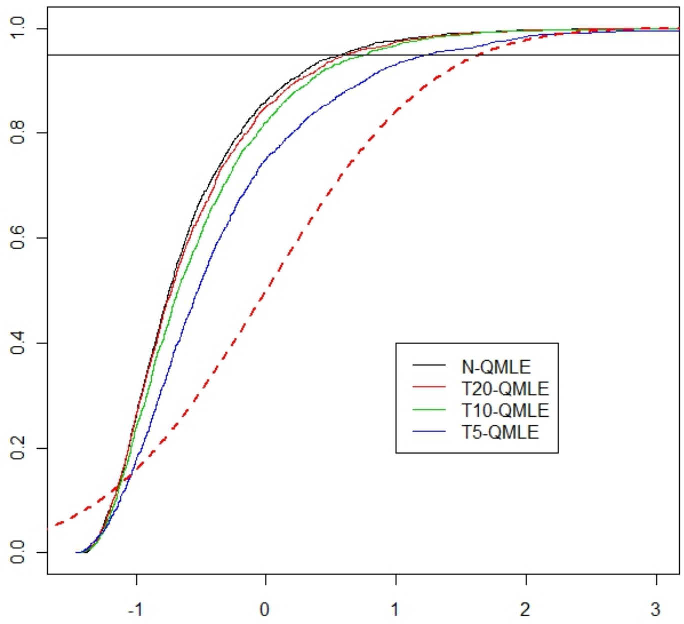
n=500, $\boldsymbol{\theta}_1 = (1, .1, .8)$ $\boldsymbol{\theta}_2 = (1, .2, .6)$



↩

Dependence on f (f_0)

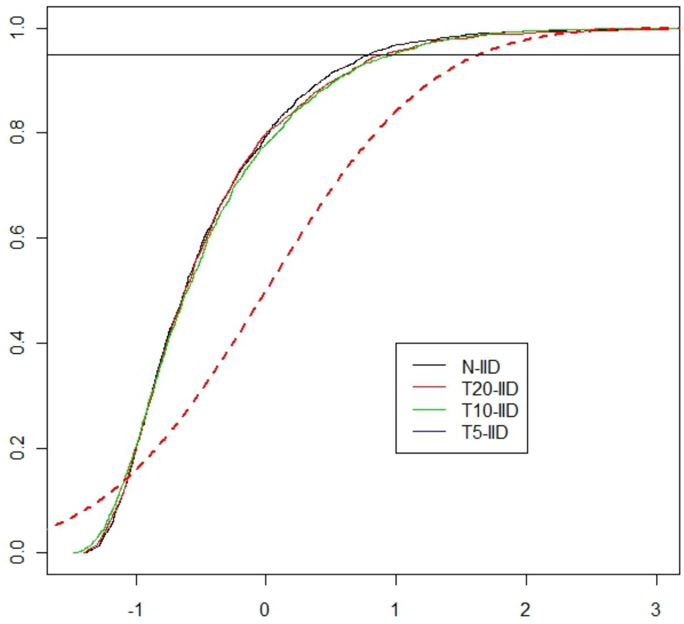
n=500



↩

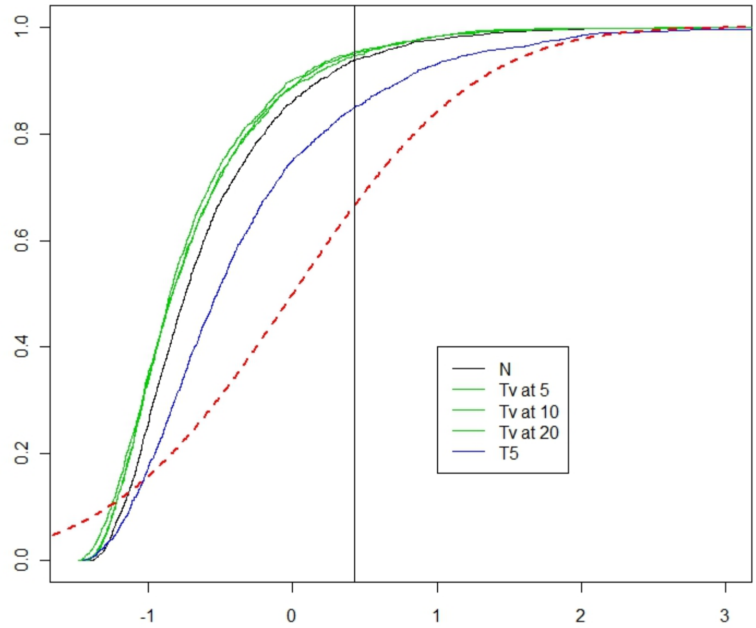
iid case

n=500



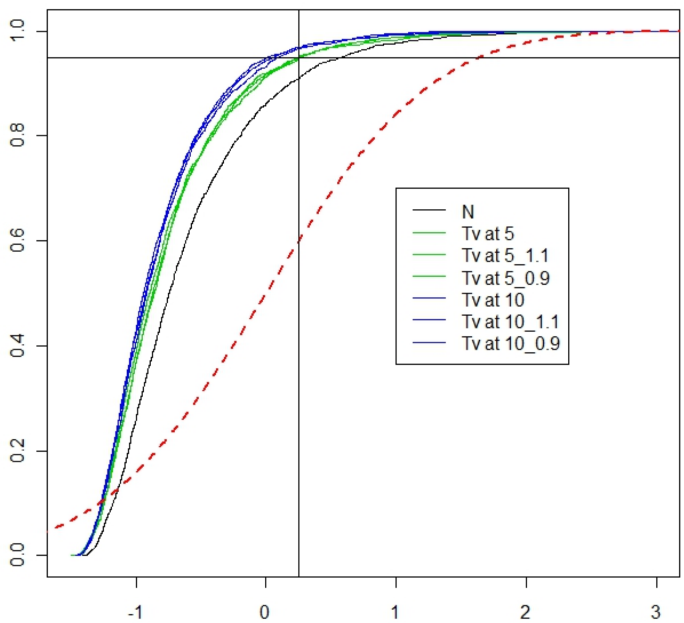
When shape parameter of T is estimated

n=500
simulated critical value for size .05: .49



When shape and skew parameters of skewed T is estimated

n=500
simulated critical value for size .05: .26



Conclusion

For finite $n < 1000$):

- dependence on $\boldsymbol{\theta}$ is neglegible.
- dependence on n significant. Convergence to the normal is not good at n=1000.
- dependence on f (f_0) is significant
- search for better choice on h.
- the test is conservative for all f considerd.

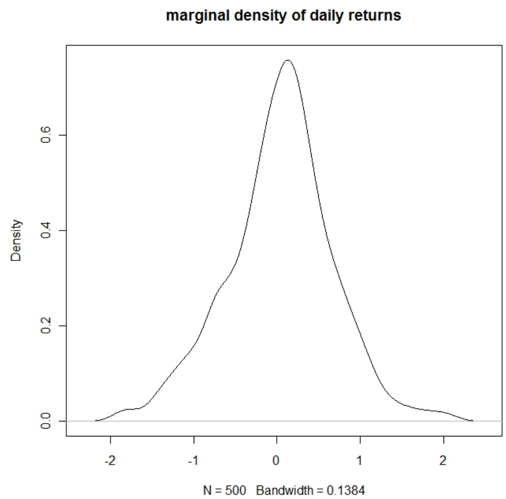
4

Application to S&P 500 Daily Return

S&P 500

Daily return from Dec 23 2004 to Dec 15 2006

- mean 0.0395
- skewness -0.114



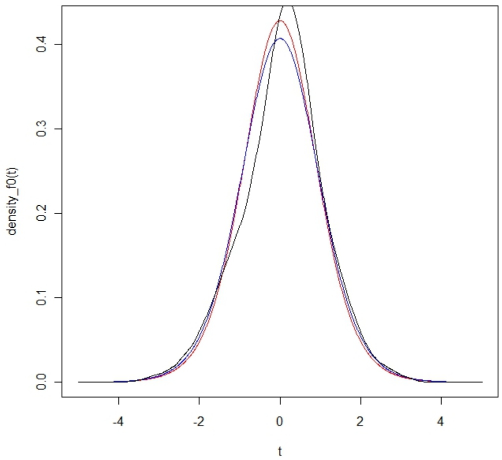
GARCH(1,1) model

with Q-MLE estimator

- $n = 500$
- μ 0.053804
- $\hat{\omega}$ 0.023656
- $\hat{\alpha}_1$ 0.060081
- $\hat{\beta}_1$ 0.881231
- Ljung-Box Test on residual for 10, 15, 20 lags : p-value $> .75$
- McLeod-Li Test on squared residual for 10, 15, 20 lags: p-value $> .43$

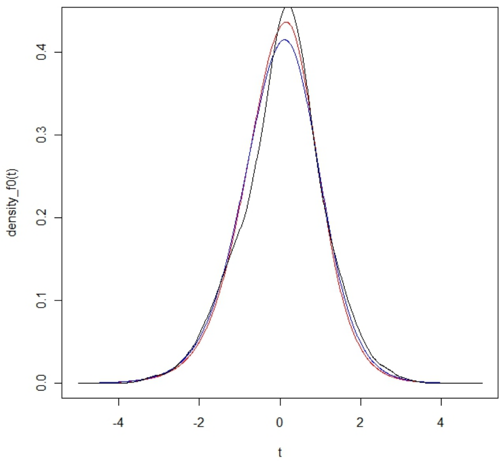
null: T(12)

- test stat $= 2.527 > 1.96$
- null is rejected



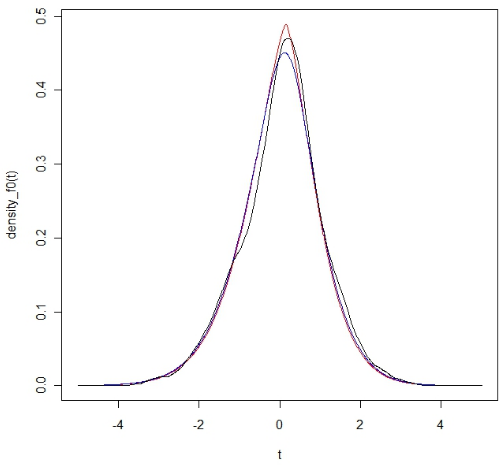
Skewed T

- MLE: shape=11.2 skew=0.902975
- simulated critical value for size .05: .49
- test stat = .559 > .49



Skewed GED

- MLE: shape=1.451578 skew=0.901739
- test stat = -0.305
- Skewed GED should be chosen over Skewed T distribution



↩

Parameter Estimates with Skewed GED MLE

- $\hat{\mu} = 0.05200$

- $\hat{\omega} = 0.01874$

- $\hat{\alpha} = 0.06491$

- $\hat{\beta}1 = 0.88945$

SD of sGED vs sSTD

Estimating parameter $\boldsymbol{\theta} = \{.02, .06, .88\}$ when f is sGED.

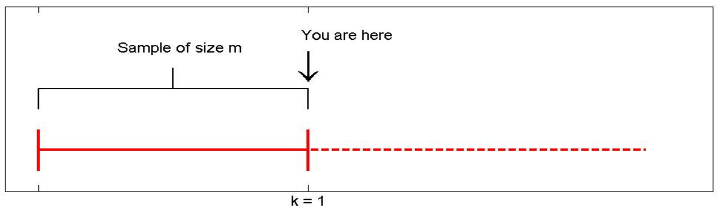
SD of estimates	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
Correct MLE (sGED)	0.0286	0.0237	0.0845
Wrong MLE(skew T)	0.0374	0.0253	0.1028
Increase	30%	7%	22%

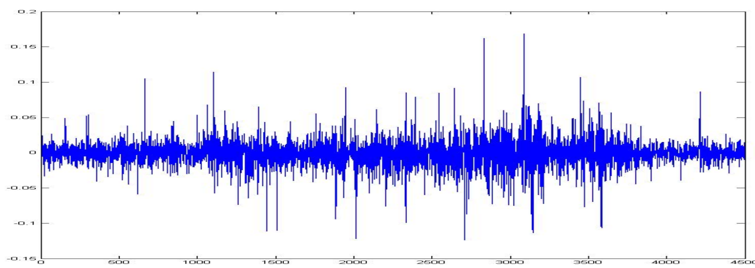
5

Sequential Monitoring of GARCH parameters

Problem

- Have an observation of size m.
- Confident that the sample is a GARCH(1,1) with a parameter $\boldsymbol{\theta} = (\omega, \alpha, \beta)$.
- Do incoming data have same parameter?





sequential monitoring

- One-shot test can only detect change-point within past observations .
- One-shot test cannot be used repeatedly.
- We have to use a sequential change-point monitoring.

$$H_0: \text{ There is no change in the parameter } \boldsymbol{\theta} \text{ for all } k \geq 1$$
$$H_A: \ \boldsymbol{\theta} \text{ changes to } \boldsymbol{\theta}^* \text{ at } k = k^* \geq 1 \ .$$

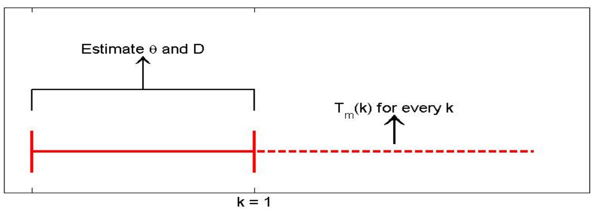
The sequential monitoring scheme

- Berkes, Gombay, Horváth, and Kokoszka (2004)

- Get $\hat{\boldsymbol{\theta}}_m, \hat{\boldsymbol{D}}_m$ from the initial sample using Q-MLE.

- For each incoming observation, calculate the monitoring random variable $T_m(k)$.

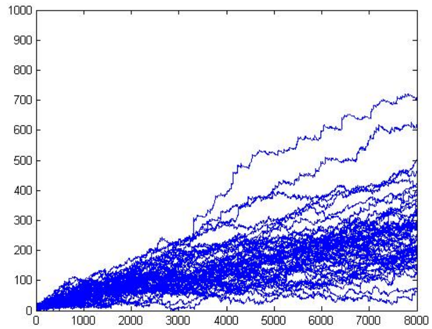
- Under H_0, $T_m(k)$ tends to stay small. Under H_A, $T_m(k)$ tends to infinity. We stop when $T_m(k)$ is 'big enough'.



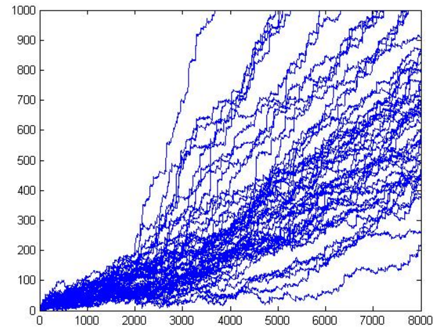
How $T_m(k)$ looks under:

50 replication with $m = 2000$

$$T_m(k) = \sum_{i=m+1}^{m+k} \hat{\ell}'_i(\hat{\boldsymbol{\theta}}_m)\hat{\boldsymbol{D}}_m^{-1/2}$$



H_0



H_A

Asymptotically Controlled Level

- Stopping time:

$$k_m(b) = \min\left\{k : T_m(k) > m^{1/2}\Big(1+\frac{k}{m}\Big)b\Big(\frac{k}{m}\Big)\right\}.$$

- Because of the convergence, if we choose $b(\cdot) = b_0$,

$$\alpha_m = P_{H_0}\{k_m(b) < \infty\} \xrightarrow{D} P\left\{\max_{1\le i\le 3}\sup_{0<s<1}|W_i(s)| > b_0\right\},$$

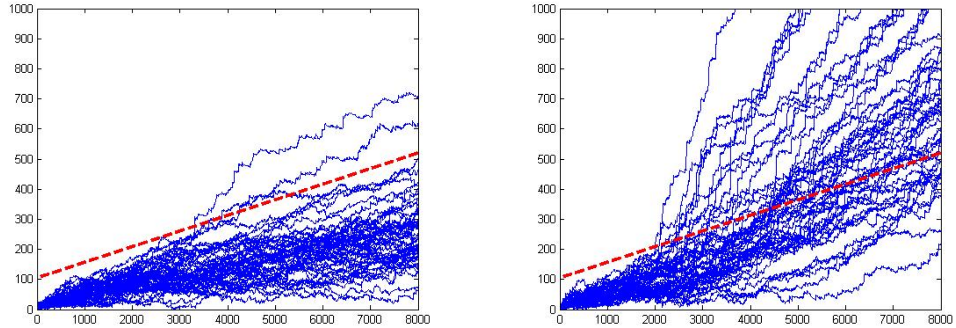
  as $m \to \infty$.

- Right hand side can be calculated by formula. By choosing the value of b_0, asymptotic level of the test can be adjusted to any value.

$T_m(k)$ with asymptotic boundary function

50 replication with $m = 2000$
$b(\cdot) = b_0 = 2.38$ (corresponding to asymptotic size of 0.1)

H_0 $\qquad\qquad$ H_A



- Q: what value of m is big enough to use the asymptotic result?

Simulation study

- Choose $b_0 = 1$.

- Want to see the convergence

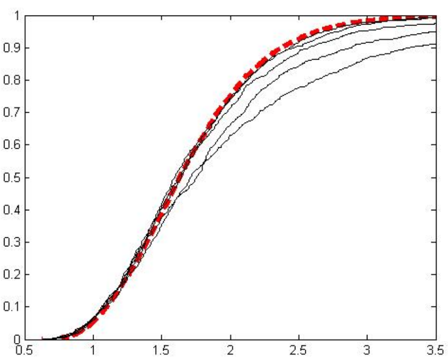
$$\sup_{1 \le k < 10^6} \frac{T_m(k)}{m^{1/2}(1+\frac{k}{m})} \quad \xrightarrow{D} \quad \max_{1 \le i \le 3} \sup_{0<s<1} |W_i(s)|.$$

  as $m \to \infty$.

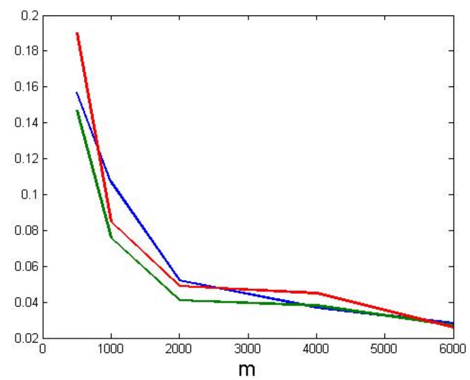
- For a particular θ, simulate with $m = 500$, 1000, 2000, 4000, 6000.

CDF comparison

$\theta = (.5, .5, .2)$



Sup of the difference



6

When m is finite

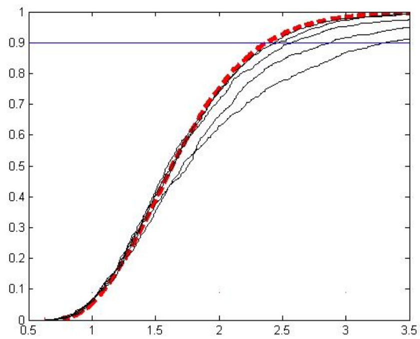
Remarks

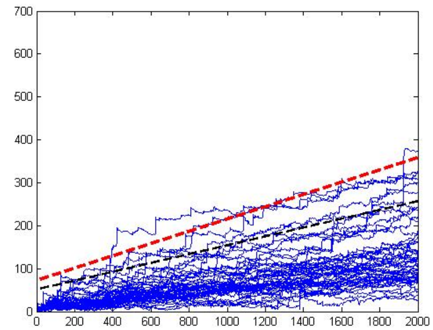
- When m is larger than 2000, (asymptotic size - actual size) is less than 5%.

- What should we do when m is less than 2000?

Option 1

- Pick b_0 accordingly.





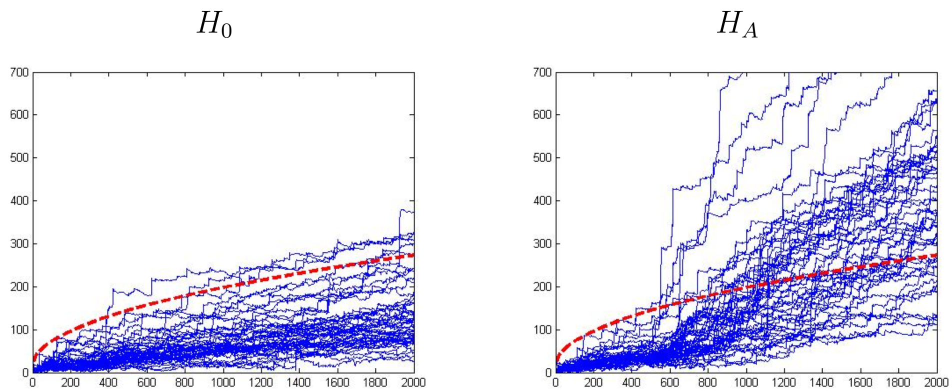
Option 2

- Construct new boundary functions $B_{\boldsymbol{\theta},m}(k)$ by simulation.
- Need to truncate the monitoring scheme.
- Modify stopping time as:

$$k_m^*(b) = \min\left\{k : T_m(k) > B_{\boldsymbol{\theta},m}(k),\ \ k \leq 4m\right\}.$$

i.e. Stops at $k = 4m$ if there is no signal.

Sample Pictures



On Truncation of the scheme

- Can use a simulation to construct boundary functions.
- Changes that occur at later k could go undetected.
- Can be used as a 'guideline' for one-shot type test.

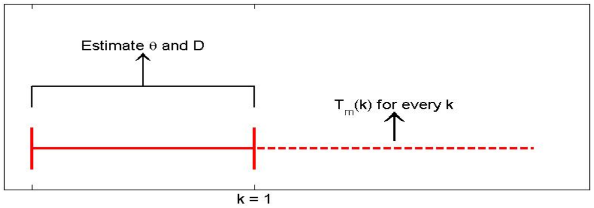
Table of Results for $m = 1000$

$B_{\boldsymbol{\theta},m}(k) = A_0 + A_1\sqrt{k}$

.8	(36.3,5.4)	-	-	-	-	-	-	-
.7	(35 ,5.7)	(5.5,5.6)	-	-	-	-	-	-
.6	(41.4,5.4)	(7.1,5.7)	(9.3,5.5)	-	-	-	-	-
.5	(21.2,5.5)	(13.6,5.7)	(1.3,5.6)	(3.3,5.8)	-	-	-	-
.4	(5.1 ,5.5)	(14.3,5.8)	(3 ,5.7)	(3.4,5.8)	(3.3,5.7)	-	-	-
.3	(2.5 ,5.1)	(3.8,5.6)	(8.5,5.7)	(0.4,5.7)	(5.5,5.7)	(1.6,5.7)	-	-
.2	(1.7,5.1)	(0.8,5.1)	(2.1,5.5)	(0.8,5.7)	(3.4,5.6)	(0 ,5.8)	(2,5.9)	-
.1	(8.2 ,4.7)	(7.7,4.7)	(14.7,4.7)	(21.8,4.7)	(16.8,5)	(17.7,5.1)	(23.7,5.3)	(28.3,5.5)
α	.1	.2	.3	.4	.5	.6	.7	.8
	β							

Empirical Power of the Modified Scheme

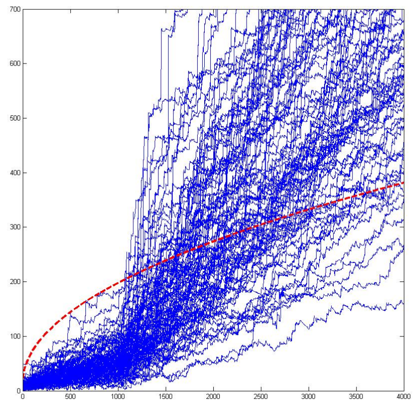
- $m = 300, 1000$
- $\boldsymbol{\theta} = (.5, .4, .2)$ changes to $\boldsymbol{\theta}^* = (\omega^*, \alpha^*, \beta^*)$ at $k = m$.
- 5000 replications



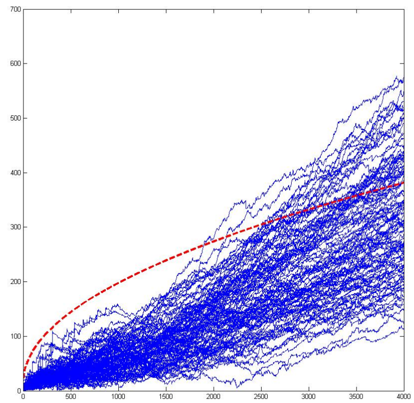
Sample Pictures

$\theta = (.5, .4, .2)$ to :

$\theta^* = (.5, .5, .3)$



$\theta^* = (.5, .3, .0)$



Results for $m = 1000$

.6	50	45	87	-	-
.5	52	20	39	98	-
.4	62	15	(10)	64	100
.3	81	28	7	23	94
.2	97	66	28	28	61
α^*	0	.1	.2	.3	.4
	β^*				

ω does not change

.7	77
.6	36
.5	(10)
.4	29
.3	100
ω^*	

(α, β) does not change

Results with perfect estimator for $m = 300, 1000$

.6	62	50	90	-	-
.5	64	24	45	98	-
.4	73	23	(14)	70	100
.3	89	41	11	28	95
.2	99	79	40	36	69
α^*	0	.1	.2	.3	.4
	β^*				

.6	100	97	100	-	-
.5	100	70	84	100	-
.4	100	87	(8)	99	100
.3	100	99	56	67	100
.2	100	100	100	98	100
α^*	0	.1	.2	.3	.4
	β^*				

.7	82
.6	44
.5	(11)
.4	40
.3	99
ω^*	

.7	100
.6	89
.5	(11)
.4	98
.3	100
ω^*	

$m = 300$ $\qquad$ $m = 1000$

7

Future Projects

↩

Future Project

Goodness-of-Fit test for error density

- Limit theorem requires $nh^5 \to \infty$. For iid case, $h \approx n^{-1/5}$ is the optimal bandwidth. What is the optimal Bandwidth for GARCH case?
- GARCH model with dynamic skewness in error distribution.
- Model assumed was pure GARCH. Can the same be said with AR-GARCH, ARMA-GARCH, etc?

Future Project

Sequential Monitoring of Parameters

- Computationally very expensive. Can we develop similar scheme with a function that is simpler to compute.
- For finite sample size, truncated version is desireble. Can the modified boundary function be independent of θ?

Thank you!