Statistics 5350/7110
Forecasting

Lecture 26
GARCH Models

Professor Robert Stine

Preliminaries

- Questions?

- Final exam
  - Similar to Midterm: short answer, multiple choice
  - One week from tomorrow: Friday, December 13
  - JMHH F45, 3-5 pm
  - Canvas announcement gives details if you don’t take the regularly scheduled final exam.

- Quick review
  - Stock returns
  - Variance changes, and volatility can be modeled
  - ARCH models

Today's Topics

Textbook §8.1

- Extending ARCH models
  - One-step ahead variance
    - ARMA models: Constant one-step ahead forecast variance
    - ARCH models: Allow that variance to change
  - Predictions more accurate in periods of less volatility

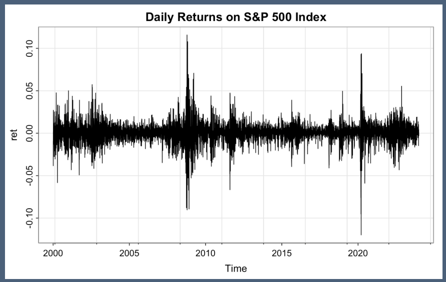
- GARCH models
  - Expand dependence from autoregressive to autoregressive, moving average
  - ARCH is to GARCH as AR is to ARMA

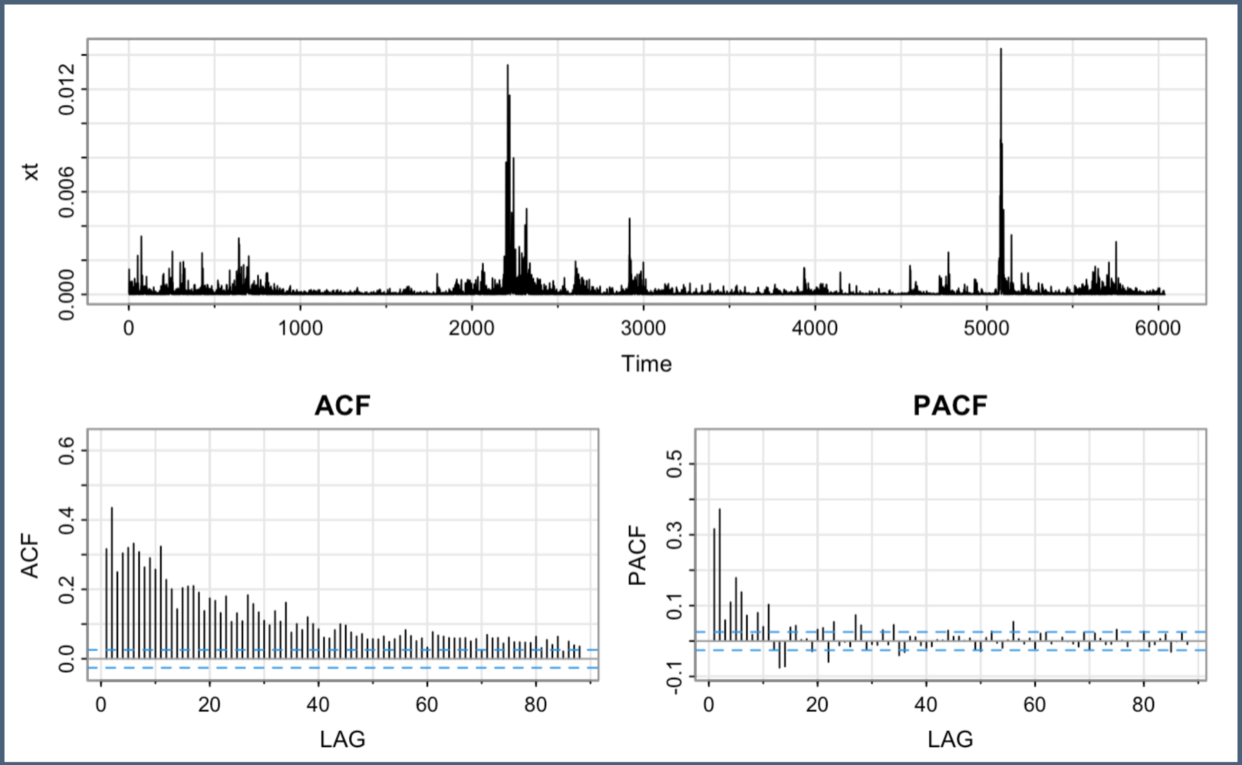
- Alternative models
  - Different target distribution from the normal: Asymmetric distributions such as exponential
  - Different model for how past returns influence volatilities

ARCH Model

Volatility

- Dependence in the squares of the returns
  - Unlike returns, substantial autocorrelation
  - Concentrated in ≈ first 10-15 lags





ARCH Model

- Autoregressive, conditional heteroscedasticity
  - Two-equation model: one equation for observation and one for hidden state
  - Proposed by Engle (1982) in study of UK inflation
- ARCH(1)
  - Gaussian noise process (independent of other terms) $\epsilon_t \sim N(0,1)$
  - Observed time series $r_t = \sigma_t \epsilon_t$ Equation 8.2
  - "Hidden" variation process (not directly observed) $\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$ Equation 8.3
  - Combine and arrange equations

$$r_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \underbrace{\sigma_t^2(\epsilon_t^2 - 1)}_{v_t}$$ Equation 8.4

  where v_t is white noise.
- Estimation
  - Maximum likelihood is essentially weighted LS, noting that $r_t \mid r_{t-1} \sim N\left(0,\, \alpha_0 + \alpha_1 r_{t-1}^2\right)$

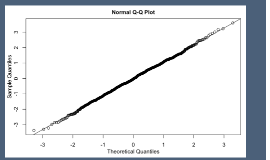
Fat tails?

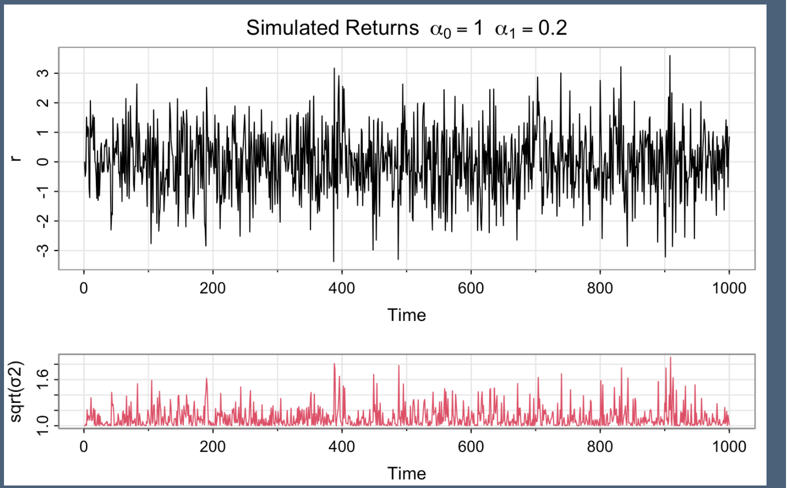
ARCH(1) Example

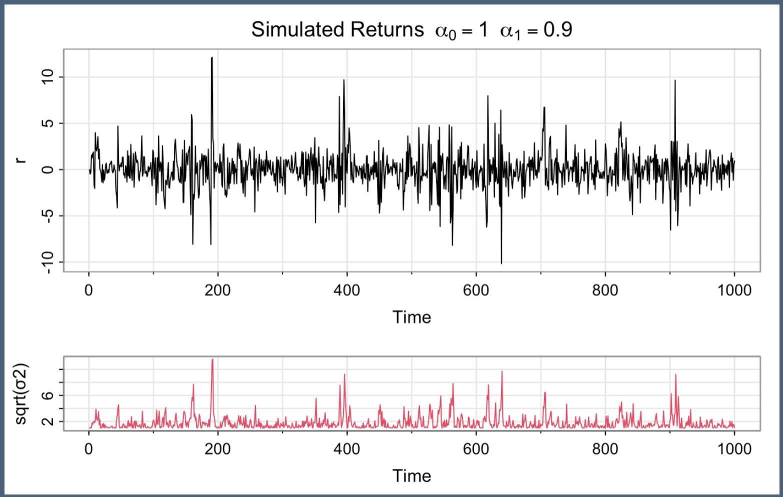
- Simulate process
  - How does estimation perform when model is correct
  - Simulate Gaussian noise process (independent of other terms) $\epsilon_t \sim N(0,1)$
  - Recursively generate the observed time series

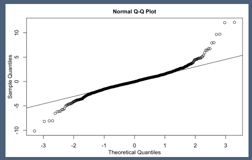
$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 \quad \text{followed by} \quad r_t = \sigma_t \epsilon_t$$

- Question
  - Do these simulated returns resemble returns on the S&P 500?









ARCH(1) Estimates

- Estimation
  - Can we recover the coefficients used in the simulation?
- Weighted least squares
  - Supply weights to the `dylm` function
  - Regress r_t^2 on it's lags; use 1/fit as weights

`weights`	an optional vector of weights to be used in the fitting process. If specified, weighted least squares is used with weights `weights` (that is, minimizing `sum(w*e^2)`); otherwise ordinary least squares is used.

WLS iterations

	α0	α1
Fit 0	1.8798	0.5423
Fit 1	1.2496	0.6958
Fit 2	1.2001	0.7079
Fit 3	1.1962	0.7088
True	1.0000	0.9000

```
Coefficients:
            Estimate Std. Error
(Intercept)  1.19616    0.13497
L(sqr, 1)    0.70883    0.04312
```

WLS underestimates coefficient in the variance equation

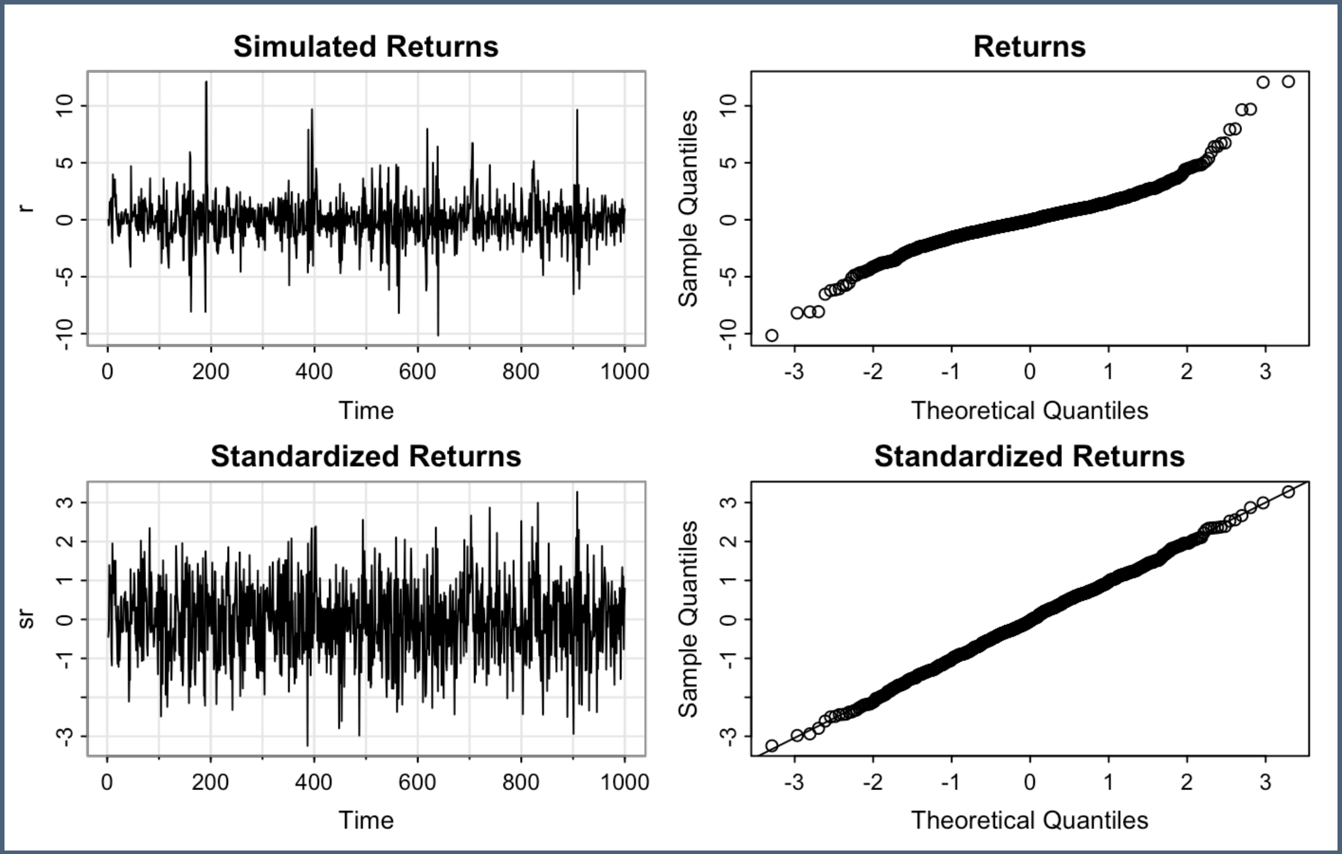
```
Error Analysis:
        Estimate  Std. Error
mu      -0.01808     0.03561
omega    1.05103     0.08680
alpha1   0.82944     0.07834
```

fGarch package does a better job for this example

fGarch includes many supplemental diagnostics as well.

Standardized Results

- Compare before/after



Using fit from WLS

ARCH(1) Estimates: S&P 500

- Fit model to daily returns
  - Use fGarch R package for ML estimates
    Highly significant coefficients
  - Diagnostics indicate problems

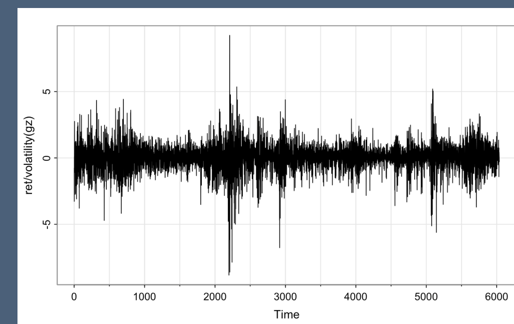
Error Analysis:

	Estimate	Std. Error	t value
mu	5.856e-04	1.326e-04	4.416
omega	9.644e-05	2.452e-06	39.331
alpha1	3.980e-01	2.817e-02	14.130

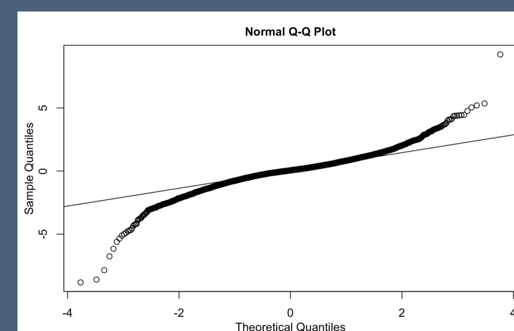
Omega = α_0

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	11917.86564	0.000000000
Shapiro-Wilk Test	R	W	NA	NA
Ljung-Box Test	R	Q(10)	19.25891	0.037094782
Ljung-Box Test	R	Q(15)	35.37306	0.002175275
Ljung-Box Test	R	Q(20)	50.42186	0.000192673
Ljung-Box Test	R^2	Q(10)	1714.96031	0.000000000
Ljung-Box Test	R^2	Q(15)	2193.77782	0.000000000
Ljung-Box Test	R^2	Q(20)	2560.64958	0.000000000



Plots of standardized residuals



- How to improve this model?
  - Better mean function?
  - Different volatility model?
  - Normality assumption?

ARCH(p) Model

Extensions of ARCH(1) Model

- Mean term
  - Allows returns to have a non-zero mean or regression trend

$$r_t = \mu_t + \sigma_t \epsilon_t$$

Equation 8.8

where the mean might be, for example, a regression

$$\mu_t = \beta_0 + \beta_1 X_{t,1} + \beta_2 X_{t,2}$$

or perhaps an ARMA model

$$\mu_t = \phi_0 + \phi_1 r_{t-1} + \theta_1 w_{t-1} + w_t$$

- ARCH(p)
  - Expand expression for conditional variance to have more lags

Equation 8.7

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \cdots + \alpha_q r_{t-p}^2$$

  - Allows wider range of behavior for volatility, such as more concentrated bursts

Constraints

- Assume ARCH(2)
  - Mean of return is 0 (assuming $\mu_t = 0$) $\quad E\, r_t = E(\sigma_t \epsilon_t) = 0$
  - Conditional variance

$$\text{Var}(r_t \,|\, r_{t-1}, r_{t-2}) = \sigma_t^2$$

Remember that $\sigma_t^2 \in \mathscr{F}_{t-1}$

  - Marginal variance

$$\text{Var}(r_t) = E\, r_t^2 - (E\, r_t)^2 = E\, r_t^2$$

which is (iterated expectations)

$$E\, r_t^2 = E\left(E\,(r_t^2 \,|\, r_{t-1}, r_{t-2})\right) = E\left(\alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2\right) = \alpha_0 + \alpha_1 E(r_{t-1}^2) + \alpha_2 E(r_{t-2}^2)$$

  - Assuming variance stationary, then

$$\text{Var}(r_t) = \alpha_0/(1 - \alpha_1 - \alpha_2)$$

- Constraints
  - Guarantee variance is not negative $\quad 0 < \alpha_0,\ 0 \le \alpha_j \quad (0 < j)$
  - Sum $\quad \alpha_1 + \alpha_2 < 1$

If model is not stationary then what are you doing using averages to estimate parameters?

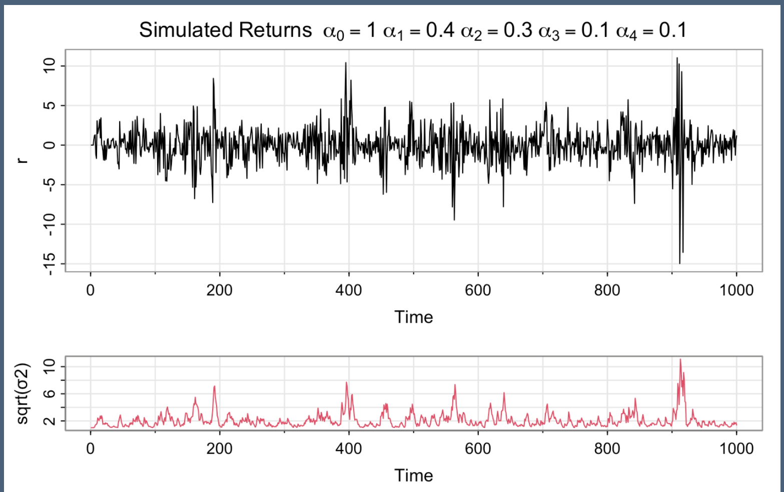
ARCH(4) Simulated Examples

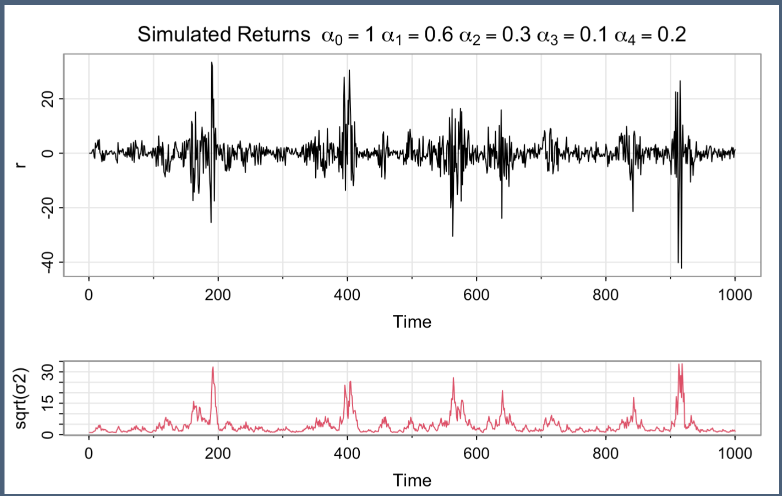
- Simulated data
  - Simulate Gaussian noise process (independent of other terms) $\epsilon_t \sim N(0,1)$
  - Recursively generate the observed time series

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \alpha_3 r_{t-3}^2 + \alpha_4 r_{t-4}^2 \quad \text{followed by} \quad r_t = \sigma_t \epsilon_t$$

- Fidelity to returns
  - Observed simulation able to mimic volatility of market return

Eeek, these parameters are not appropriate!



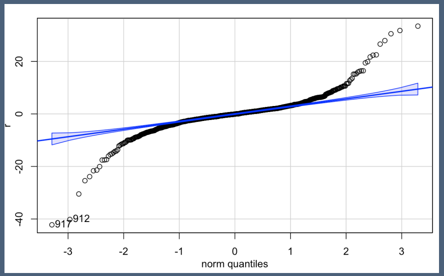


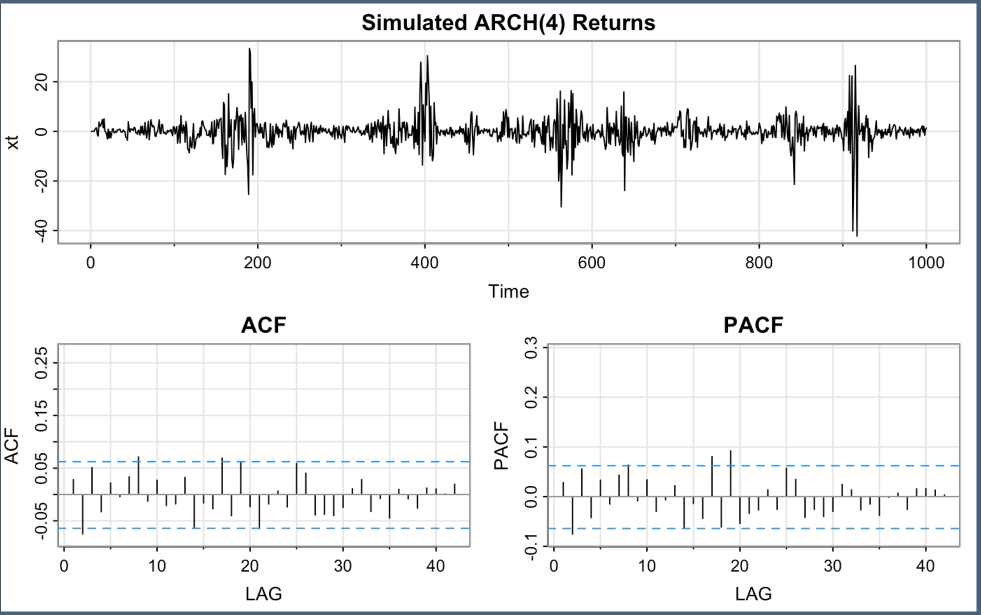
Better match to returns on S&P

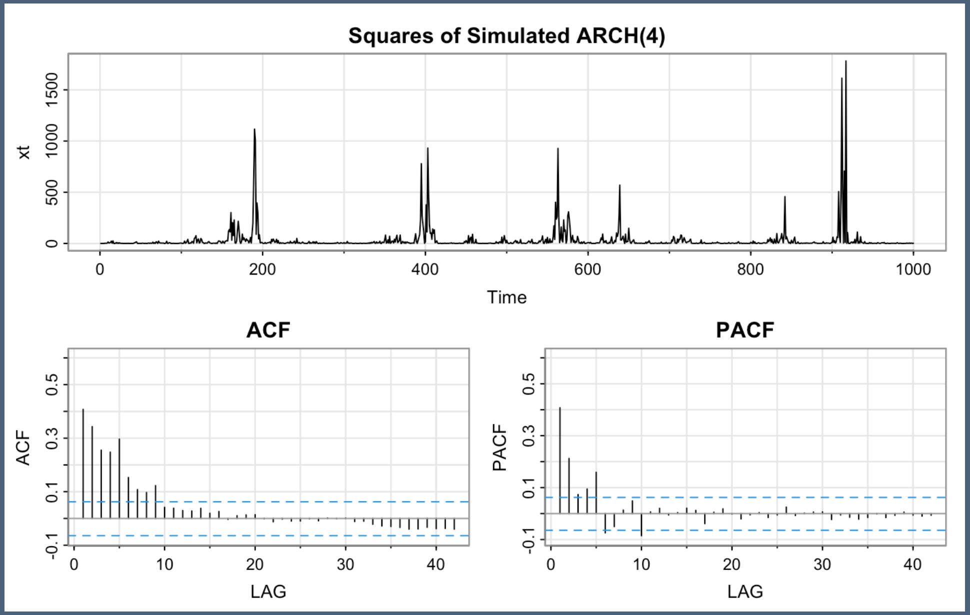
Appearance changes depending on the random seed.

ARCH(4) Example

- Properties of simulated returns
  - Returns are more fat tailed than ARCH(1)
  - Returns lack much dependence
  - Squared returns are autocorrelated







ARCH(4) Simulated Estimates

- Estimation
  - Can we recover the coefficients used in the simulation?
  - Weighted least squares again effective
- Weighted least squares
  - Supply weights to the `dylm` function
  - Regress r_t^2 on it's lags; use 1/fit as weights

`weights`	an optional vector of weights to be used in the fitting process. If specified, weighted least squares is used with weights `weights` (that is, minimizing `sum(w*e^2)`); otherwise ordinary least squares is used.

	α_0	α_1	α_2	α_3	α_4
True	10.2749	0.3097	0.0704	0.0111	0.2385
Fit 0	3.5547	0.4675	0.1880	0.0183	0.1982
Fit 1	2.4927	0.4734	0.2114	0.0385	0.1871
Fit 2	2.2747	0.4759	0.2130	0.0429	0.1865
Fit 3	2.2298	0.4764	0.2132	0.0438	0.1865
Fit 4	1.0000	0.6000	0.3000	0.1000	0.2000

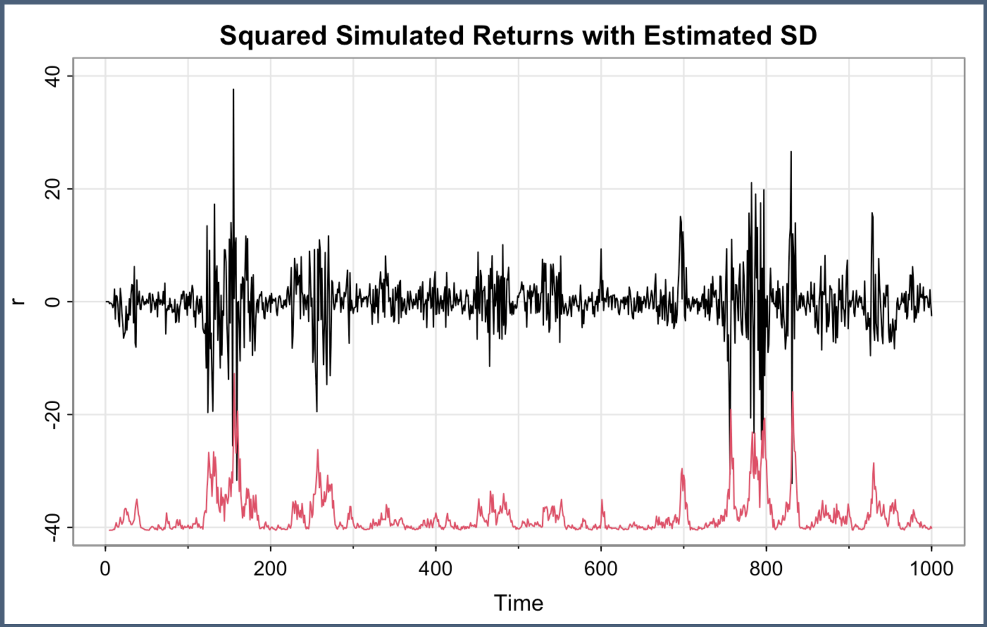
Coefficients:

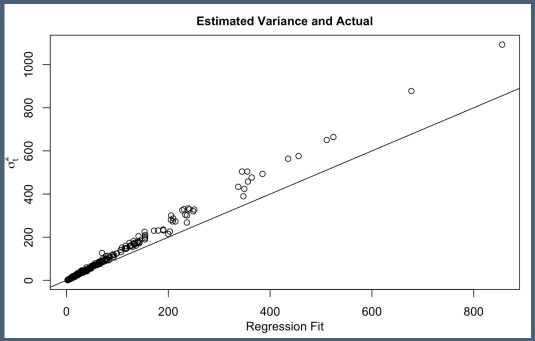
	Estimate	Std. Error	t value
(Intercept)	2.22983	0.63843	3.493
L(sqr, 1:4)1	0.47639	0.04290	11.104
L(sqr, 1:4)2	0.21319	0.03690	5.777
L(sqr, 1:4)3	0.04380	0.02971	1.474
L(sqr, 1:4)4	0.18648	0.03344	5.577

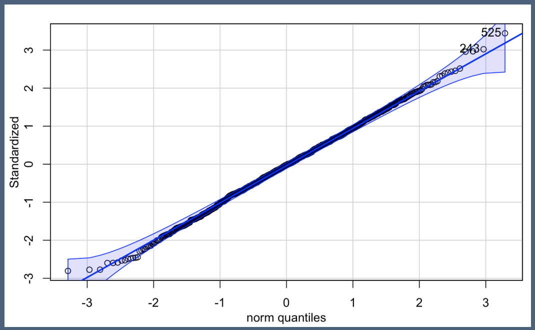
R-code: garchFit again does better!

ARCH(4) Example

- Fitted values from regression
  - Estimates of conditional variance σ_t^2
  - Much more visually similar to bursts of volatility in the SP500 returns
  - WLS underestimates large variances but standardized residuals look normal (qqplot)





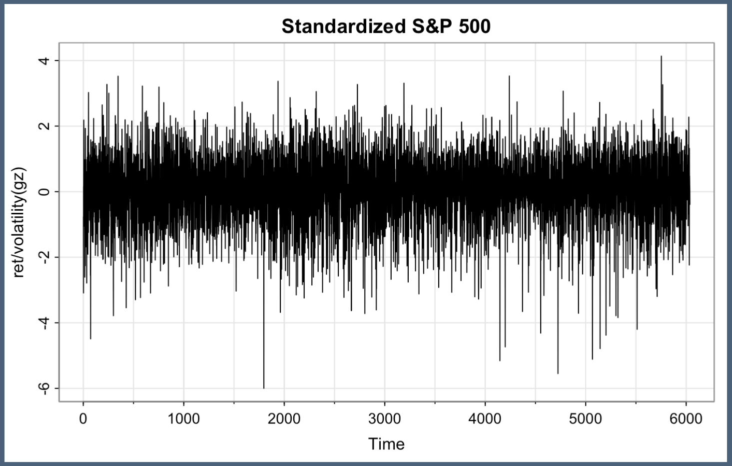


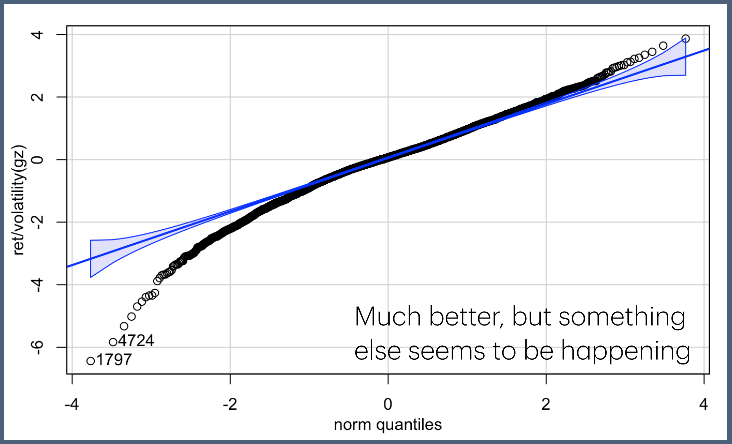
ARCH(p) for S&P 500 Returns

- Fit ARCH(12) to S&P 500 returns
  - More dependence than 4 lags: Significant out to 11 lags
  - Lots of AR terms motivates GARCH models

Error Analysis:

	Estimate	Std. Error	t value
mu	6.716e-04	1.022e-04	6.573
omega	1.917e-05	1.436e-06	13.354
alpha1	8.174e-02	1.403e-02	5.827
alpha2	1.510e-01	1.730e-02	8.729
alpha3	1.056e-01	1.666e-02	6.336
alpha4	1.205e-01	1.731e-02	6.961
alpha5	7.380e-02	1.477e-02	4.997
alpha6	6.058e-02	1.501e-02	4.037
alpha7	4.960e-02	1.402e-02	3.537
alpha8	8.255e-02	1.576e-02	5.239
alpha9	4.875e-02	1.424e-02	3.422
alpha10	5.531e-02	1.428e-02	3.874
alpha11	4.739e-02	1.371e-02	3.457
alpha12	5.162e-03	1.147e-02	0.450





GARCH Models

GARCH Model

Textbook, Example 8.2

- Generalization of ARCH model
  - Analogous to addition of moving average terms to an autoregression
  - Proposed by Bollerslev (1986) to reduce need for many terms in declining lag structure
- GARCH(1,1)
  - Gaussian noise process (independent of other terms) $\epsilon_t \sim N(0,1)$
  - Observed time series $r_t = \sigma_t \epsilon_t$ Equation 8.2
  - "Hidden" variation process (not directly observed) $\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ Equation 8.9
  - Combine and arrange equations

$$r_t^2 = \alpha_0 + (\alpha_1 + \beta_1) r_{t-1}^2 + v_t - \beta_1 v_{t-1}$$

  where $v_t = \sigma_t^2(\epsilon_t - 1)$ plays the role of white noise (albeit heteroscedastic)
- Estimation
  - WLS not so easy since regressors are not directly observed (as in ARMA(p,q) models with 0 < q)
  - Rely on maximum likelihood

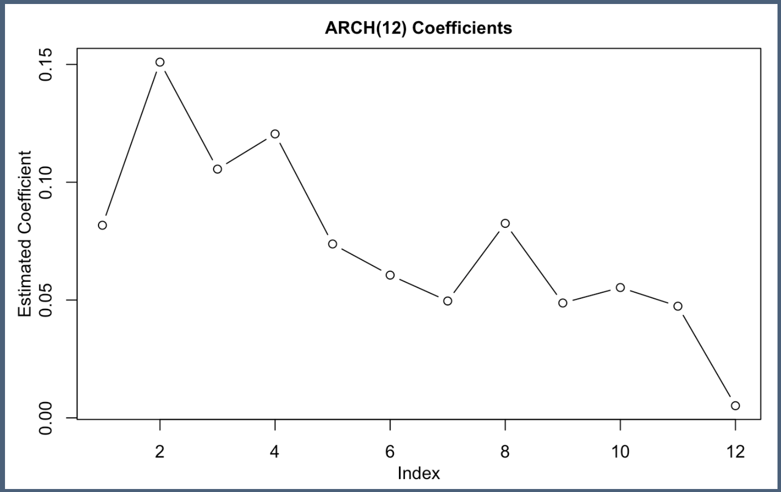
GARCH(1,1) for S&P 500 Returns

- Simplifies estimates
  - AR coefficients in ARCH have generally decreasing magnitude
  - Irregular decay, with seemingly random variation
  - Requires many parameter estimates (easy with n=6,000, harder if n=143 as in Bollerslev paper)

omega = α_0

Error Analysis:

	Estimate	Std. Error	t value
mu	6.716e-04	1.022e-04	6.573
omega	1.917e-05	1.436e-06	13.354
alpha1	8.174e-02	1.403e-02	5.827
alpha2	1.510e-01	1.730e-02	8.729
alpha3	1.056e-01	1.666e-02	6.336
alpha4	1.205e-01	1.731e-02	6.961
alpha5	7.380e-02	1.477e-02	4.997
alpha6	6.058e-02	1.501e-02	4.037
alpha7	4.960e-02	1.402e-02	3.537
alpha8	8.255e-02	1.576e-02	5.239
alpha9	4.875e-02	1.424e-02	3.422
alpha10	5.531e-02	1.428e-02	3.874
alpha11	4.739e-02	1.371e-02	3.457
alpha12	5.162e-03	1.147e-02	0.450



GARCH(1,1) for S&P 500 Returns

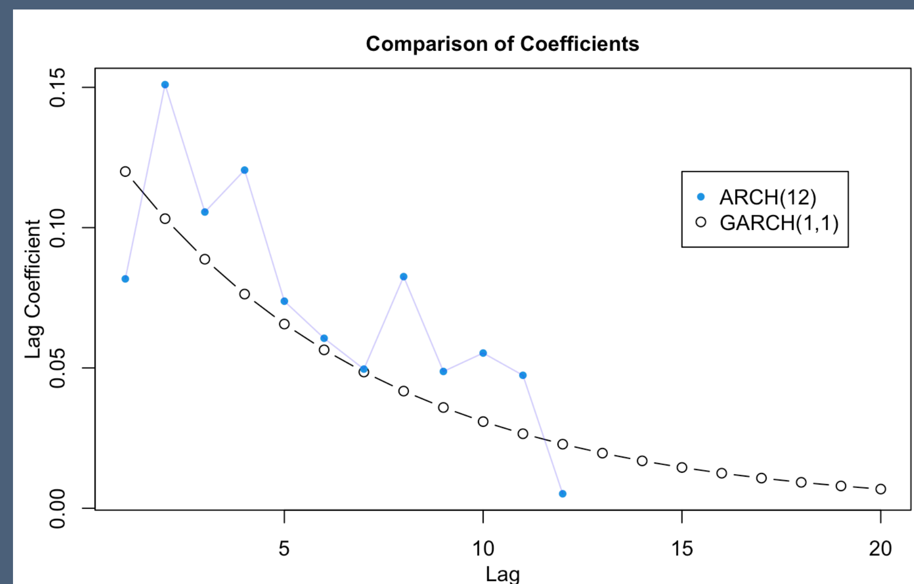
- Simplifies estimates
  - AR coefficients in ARCH have decreasing magnitude
  - ARMA coefficients simplify them, enforcing smooth weight decay

Error Analysis:

	Estimate	Std. Error	t value
mu	6.716e-04	1.022e-04	6.573
omega	1.917e-05	1.436e-06	13.354
alpha1	8.174e-02	1.403e-02	5.827
alpha2	1.510e-01	1.730e-02	8.729
alpha3	1.056e-01	1.666e-02	6.336
alpha4	1.205e-01	1.731e-02	6.961
alpha5	7.380e-02	1.477e-02	4.997
alpha6	6.058e-02		
alpha7	4.960e-02		
alpha8	8.255e-02		
alpha9	4.875e-02		
alpha10	5.531e-02		
alpha11	4.739e-02		
alpha12	5.162e-03		

Error Analysis:

	Estimate	Std. Error	t value
mu	6.397e-04	1.030e-04	6.212
omega	2.248e-06	2.855e-07	7.874
alpha1	1.211e-01	9.172e-03	13.202
beta1	8.630e-01	9.449e-03	91.333



GARCH(2,2) for S&P 500 Returns

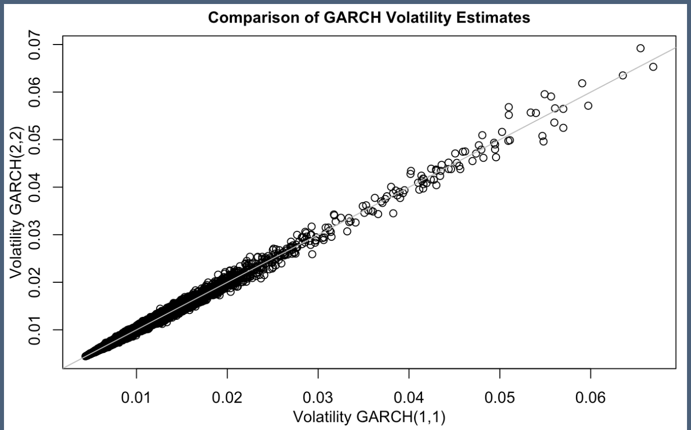
- Common practice
  - Most applications stop with the GARCH(1,1) model
  - Residual dependence diagnostics look okay

```
Ljung-Box Test     R    Q(10)   20.72505 0.02309455
Ljung-Box Test     R    Q(15)   29.49721 0.01386936
Ljung-Box Test     R    Q(20)   33.73966 0.02794068
Ljung-Box Test     R^2  Q(10)   15.03733 0.13070682
Ljung-Box Test     R^2  Q(15)   18.19571 0.25251660
Ljung-Box Test     R^2  Q(20)   19.41221 0.49519529
```

- More complex model
  - Add additional terms, moving to GARCH(2,2) model
  - Added coefficients are statistically significant
  - Max of log-likelihood improves from 19448 to 19455
    but remember that statistical significance does not imply meaningfully different

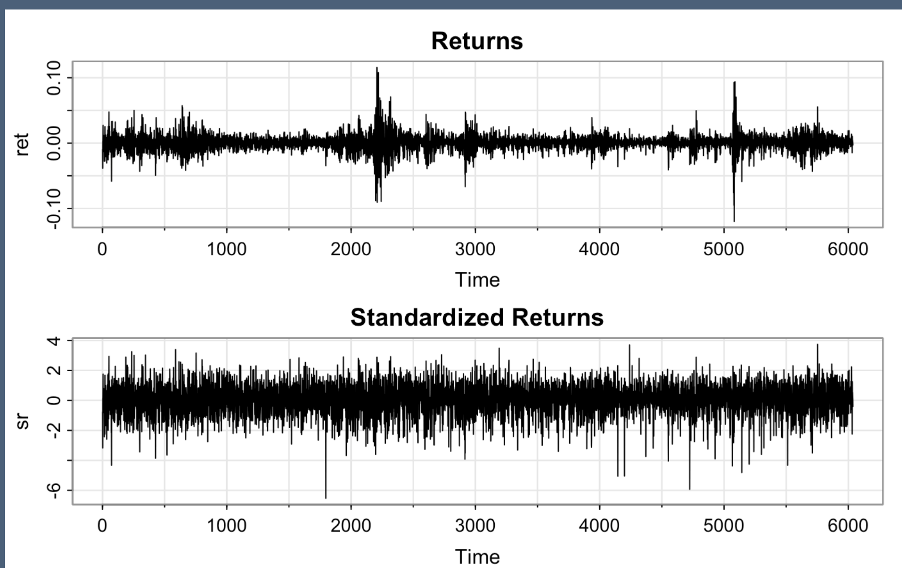
  - For example, small change to the estimated volatilities

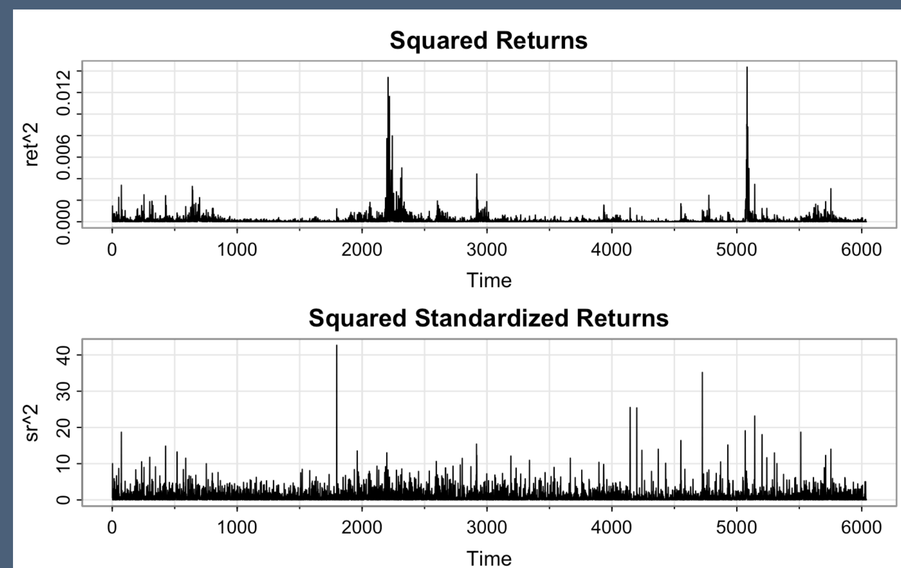
```
        Estimate  Std. Error  t value Pr(>|t|)
mu     6.442e-04   1.027e-04    6.274 3.52e-10
omega  4.086e-06   5.798e-07    7.046 1.84e-12
alpha1 8.626e-02   1.387e-02    6.220 4.98e-10
alpha2 1.259e-01   1.895e-02    6.647 3.00e-11
beta1  2.138e-01   1.808e-01    1.182 0.237149
beta2  5.448e-01   1.618e-01    3.368 0.000756
```



GARCH(1,1) Standardized Returns

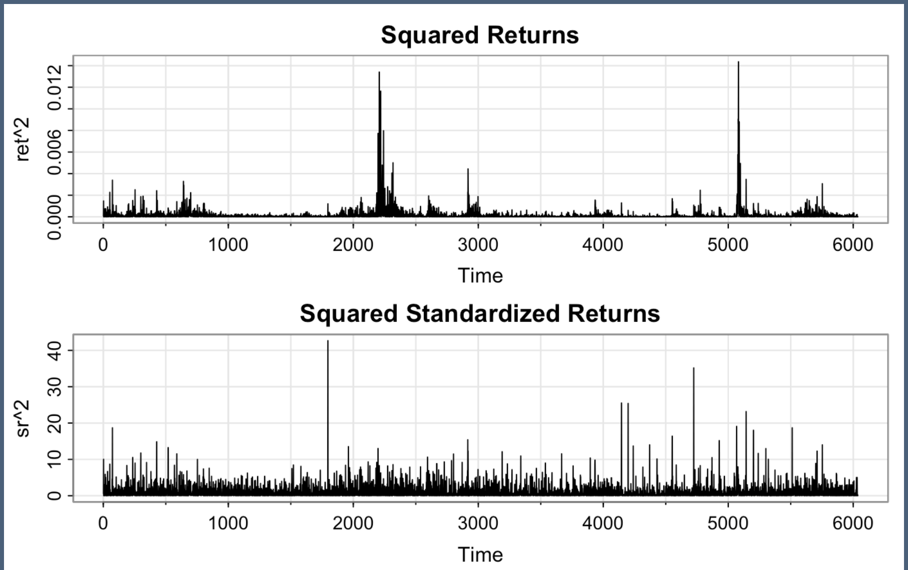
- Inspect the standardized returns for
  - Changes in volatility … ACF of squared standardized returns (automatic diagnostics)
  - Normality of distribution
- Results: Sequence plots

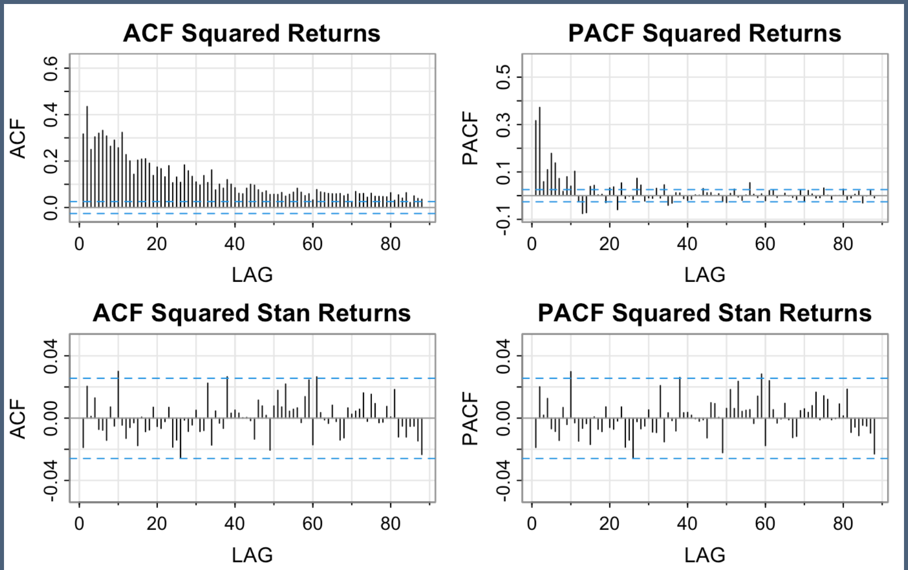




GARCH(1,1) Standardized Returns

- Inspect the standardized returns for
  - Changes in volatility … ACF of squared standardized returns (automatic diagnostics)
  - Normality of distribution
- Results: Autocorrelations

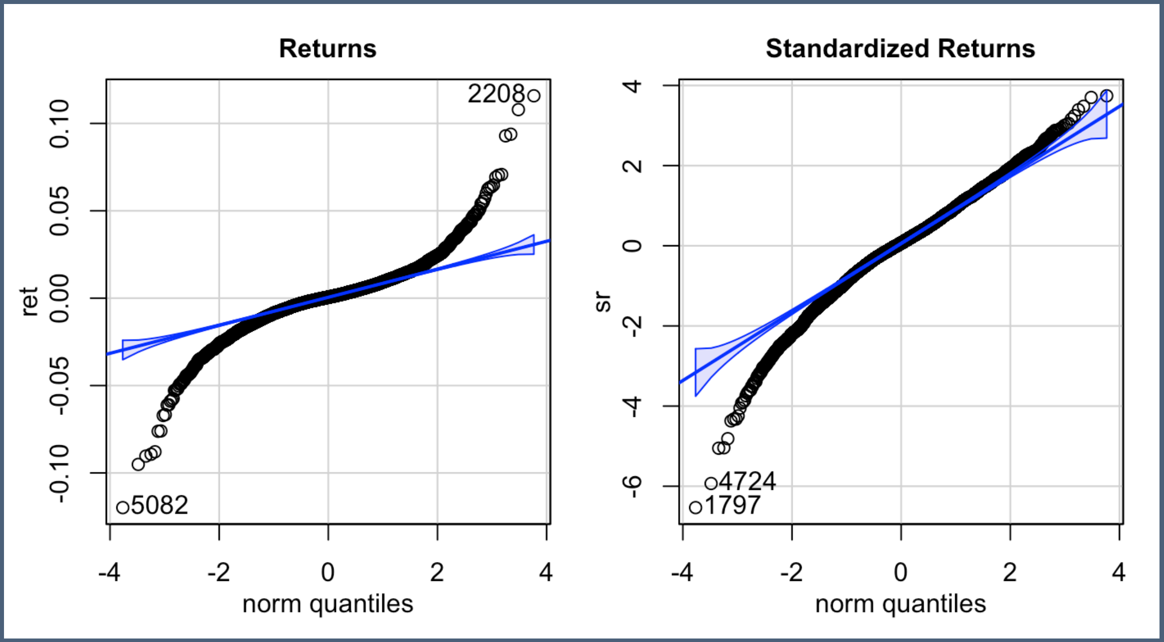




GARCH(1,1) Standardized Returns

- Inspect the standardized returns for
  - Changes in volatility … ACF of squared standardized returns (automatic diagnostics)
  - Normality of distribution

- Results: Distribution
  - Closer to normal, but…
  - Deviations from normality lie more in asymmetry rather than fat tails.



Asymmetric Power GARCH

Asymmetric Power GARCH

Textbook, Example 8.3

- Motivation
  - Distribution of standardized returns is not symmetric
  - More volatility on lower side of the distribution
- Solutions
  - Target an asymmetric error model, something other than normal
    - Exponential
    - Skew normal
  - Introduce asymmetry into the model for volatilities
- AP-GARCH
  - Keep the observation equation, but change the model for volalilities to (AP-GARCH(1,1))

$$\sigma_t^\delta = \alpha_0 + \alpha_1 \left(|r_{t-1}| - \gamma_1 r_{t-1} \right)^\delta + \beta_1 \sigma_{t-1}^\delta$$

  - Reduces to GARCH(1,1) if

$$\gamma_1 = 0, \quad \delta = 2$$

  - $0 < \gamma_1$ implies that negative shocks have larger impact than positive shocks.

AP-GARCH for S&P 500 Returns

- Estimate model
  - With estimated coefficients inserted…

$$\sigma_t^{\delta} = \omega + \alpha_1 \left(|r_{t-1}| - \gamma_1 r_{t-1} \right)^{\delta} + \beta_1 \sigma_{t-1}^{\delta}$$

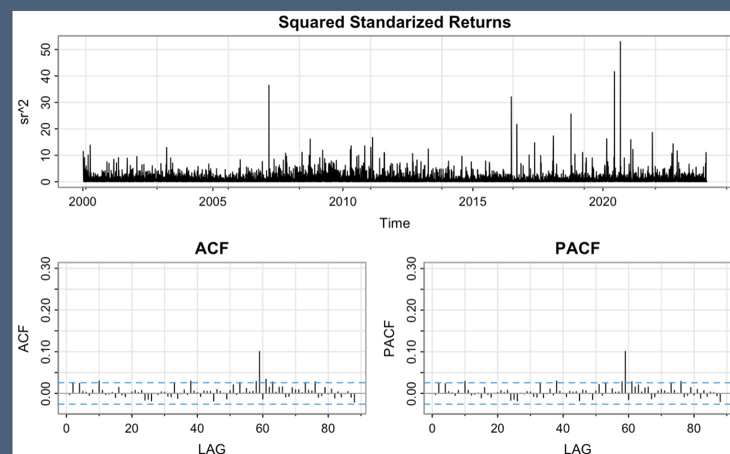
$$\sigma_t = 0.0003 + 0.09 \left(|r_{t-1}| - 0.94\, r_{t-1} \right) + 0.9\, \sigma_{t-1}$$

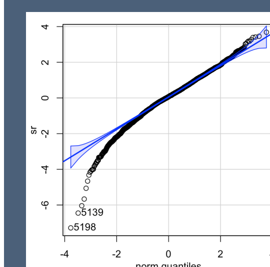
	Estimate	Std. Error	t value
mu	1.750e-04	9.248e-05	1.893
omega	3.066e-04	3.218e-05	9.526
alpha1	9.164e-02	7.860e-03	11.658
gamma1	9.382e-01	6.690e-02	14.024
beta1	9.024e-01	7.038e-03	128.219
delta	9.886e-01	8.512e-02	11.613

  - Estimated coefficients appear significant
- Residual diagnostics
  - Lots of problems, but for absence of autocorrelation in returns
  - Substantial correlation left in standardized returns, doesn't fix asymmetry…

Standardised Residuals Tests:

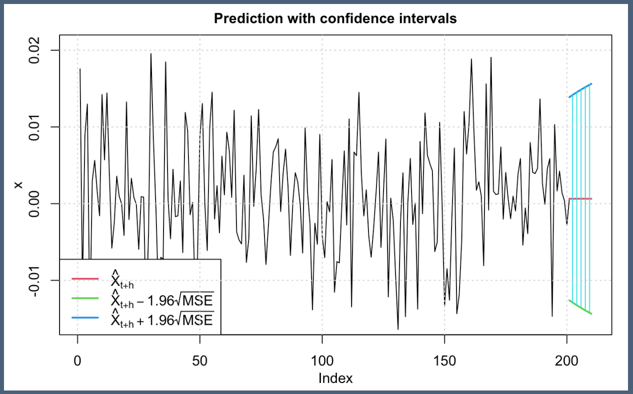
			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	561726.96973	0.0000000
Shapiro-Wilk Test	R	W	NA n is too large	NA
Ljung-Box Test	R	Q(10)	11.14533	0.3463079
Ljung-Box Test	R	Q(15)	17.81399	0.2725710
Ljung-Box Test	R	Q(20)	23.48928	0.2654129
Ljung-Box Test	R^2	Q(10)	839.51845	0.0000000
Ljung-Box Test	R^2	Q(15)	840.43692	0.0000000
Ljung-Box Test	R^2	Q(20)	840.72521	0.0000000
LM Arch Test	R	TR^2	13.18454	0.3557712





Discussion of ARCH Models

- Numerous variations of the original ARCH model
  - WE cover GARCH, AP-GARCH
  - Wikipedia lists more than 10 flavors of ARCH models
- Constraints
  - Stationarity requires elaborate restrictions on the parameters of an
- Weakly predictive
  - Graph shows predicted returns from GARCH(1,1) model
  - Mean-reversion does not anticipate large shocks to system



- Non-structural
  - Like ARMA models, GARCH models do not explain why the volatility changes … only that it does
- Why bother with them?
  - Volatility is an input to other models in Finance
  - We don’t observe volatility directly, but can use estimates from a model

What's next?

- Final class
- Review session

Come with your questions or it will be a very short class!