MM955 Financial Econometrics Coursework

Part A - Test on Theory

Question 1

(a) (1) State the relation between the single-period log-returns and the multiperiod log return.

The relation between single-period log-returns and the multiperiod log-return is that the multiperiod log-return equals the sum of the individual single-period log-returns. Mathematically, if we denote the single-period log-return for period t as:

$$r_t = \ln(P_t/P_{t-1})$$

And the multiperiod log-return over n periods as:

$$R_{0,n} = \ln(P_n/P_0)$$

Then:

$$R_{0,n} = r_1 + r_2 + \ldots + r_n = \sum_{t=1}^{n} r_t$$

This additive property is a key advantage of using logarithmic returns in financial analysis.

(a) (2) Briefly state the reasons why many financial studies involve returns, instead of prices.

For financial studies, there are three main reasons why returns are preferred over prices:

1. Returns are comparable across different assets, regardless of their price levels, allowing for meaningful cross-sectional analysis.
2. Returns have more attractive statistical properties than prices - they tend to be stationary and closer to a normal distribution, making them more suitable for statistical inference and modeling.
3. Returns directly measure the investment performance and financial gain, which is what most investors and analysts are ultimately interested in evaluating.

(a) (3) If the quarterly log-return of an asset is 2.05%, what is the corresponding quarterly relative return?

To find the quarterly relative return from the given quarterly log-return, I need to use the relationship between log-returns and relative returns.

If the quarterly log-return is 2.05%, I can express this as:

$$r_log = \ln(1 + r_{relative}) = 0.0205$$

To find the relative return, I need to solve for r_relative:

$$r_{relative} = e^{r_{log}} - 1 = e^{0.0205} - 1$$

Calculating:

$$r_{relative} = e^{0.0205} - 1 = 1.0207 - 1 = 0.0207 = 2.07\%$$

Therefore, the corresponding quarterly relative return is 2.07%.

(a) (4) If every monthly log-return of an asset within a year is 1.03%, what is the yearly log-return of the asset?

If the monthly log-return is 1.03% for every month in a year, then we can calculate the yearly log-return by using the additive property of log-returns.

Since there are 12 months in a year, and each month has the same log-return of 1.03%, the yearly log-return would be:

$$R_{year} = \sum_{t=1}^{12} r_t = 12 \times 0.0103 = 0.1236 = 12.36\%$$

Therefore, the yearly log-return of the asset is 12.36%.

(a) (5) If every quarterly relative return is 1.94% in a year, what is the yearly relative return?

The yearly relative return can be calculated using the multiplicative property of relative returns. For quarterly compounding, the relationship is:

$$(1 + r_{year}) = (1 + r_q)^4$$

Where r_q is the quarterly relative return and 4 is the number of quarters in a year. Substituting the quarterly return of 1.94%:

$$(1 + r_{year}) = (1 + 0.0194)^4 = (1.0194)^4 = 1.0798$$

Therefore:

$$r_{year} = 0.0798 = 7.98$$

The yearly relative return is 7.98%.

(b) (1) The yearly return of an asset is denoted by R. Let f be the probability density function of R. The yearly Value at Risk (VaR) with confidence level α is denoted by VaRα.

To express the Value at Risk (VaRα) in terms of the probability density function f:

$$P(R \leq -VaR_\alpha) = 1 - \alpha$$

Using the probability density function:

$$\int_{-\infty}^{-VaR_\alpha} f(r)dr = 1 - \alpha$$

This equation states that VaR_α is the value where the cumulative probability of returns below $-VaR_\alpha$ equals $1-\alpha$.

(b) (2) Suppose that R follows a normal distribution $N(0.3, \sigma^2)$. Express VaRα in terms of σ and $\Phi^{-1}(1 - \alpha)$, where $\Phi^{-1}(1 - \alpha)$ is the (1 - α) quantile of the standard normal distribution N(0, 1).

To find VaRα for $R \sim N(0.3, \sigma^2)$:

From the definition of VaRα:

$$P(R \leq -VaR_\alpha) = 1 - \alpha$$

Since R follows a normal distribution with mean 0.3 and variance σ^2, we can standardize:

$$P\left(\frac{R - 0.3}{\sigma} \leq \frac{-VaR_\alpha - 0.3}{\sigma}\right) = 1 - \alpha$$

Let $Z = (R - 0.3)/\sigma$, where $Z \sim N(0,1)$. Then:

$$P\left(Z \leq \frac{-VaR_\alpha - 0.3}{\sigma}\right) = 1 - \alpha$$

By definition of the standard normal quantile function:

$$\frac{-VaR_\alpha - 0.3}{\sigma} = \Phi^{-1}(1 - \alpha)$$

Solving for VaRα:

$$-VaR_\alpha - 0.3 = \sigma\Phi^{-1}(1 - \alpha)$$

$$-VaR_\alpha = 0.3 + \sigma\Phi^{-1}(1 - \alpha)$$

Therefore:

$$VaR_\alpha = -0.3 - \sigma\Phi^{-1}(1 - \alpha)$$

(b) (3) If it is given that R follows normal distribution $N(0.3,(0.4)^2)$, when α = 0.95, what is the value of VaRα? Explain the meaning of this value in finance. (Note that $\Phi^{-1}(0.05) = -1.64$)

To calculate VaRα when $R \sim N(0.3, (0.4)^2)$ and α = 0.95:

From the previous derivation, we know that:

$$VaR_\alpha = -0.3 - \sigma\Phi^{-1}(1 - \alpha)$$

Substituting the given values:

- α = 0.95, so 1-α = 0.05
- σ = 0.4
- $\Phi^{-1}(0.05) = -1.64$

$$VaR_{0.95} = -0.3 - 0.4 \times (-1.64)$$

$$VaR_{0.95} = -0.3 + 0.656$$

$$VaR_{0.95} = 0.356 \text{ or approximately } 0.36$$

The financial meaning of $VaR_{0.95}$ = 0.36 (or 36% if expressed as a percentage):

This value indicates that with 95% confidence, the maximum loss over a one-year period will not exceed 36% of the investment. In other words, there is only a 5% probability that the investment will lose more than 36% of its value in a year. Value at Risk (VaR) is a key risk management metric that helps financial institutions and investors quantify the potential for significant losses under normal market conditions, allowing them to set aside appropriate capital reserves and make informed investment decisions.

Question 2

(a) Let $\{\varepsilon_t\}$ be the usual white noise process of mean zero and variance σ^2. Define a random walk model where $X_t = X_{t-1} + \varepsilon_t$ with $X_0 = Y$, where Y is independent of $\{X_t, t \geq 1\}$ and follows normal distribution $N(\mu_Y, \sigma_Y^2)$.

(1) Why this model is nonstationary?

To determine whether the random walk model $X_t = X_{t-1} + \varepsilon_t$ with $X_0 = Y$ is stationary, I need to examine if its statistical properties (mean, variance, and autocorrelation) remain constant over time.

First, let's find the mean of X_t by recursive substitution:

$$X_t = X_{t-1} + \varepsilon_t$$

$$= X_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

$$= \ldots$$

$$= X_0 + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t$$

$$= Y + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t$$

Taking the expectation:

$$E[X_t] = E[Y] + E[\varepsilon_1] + E[\varepsilon_2] + \ldots + E[\varepsilon_t]$$

$$= \mu_Y + 0 + 0 + \ldots + 0$$

(since $E[\varepsilon_t] = 0$ for white noise)

$$= \mu_Y$$

Now for the variance:

$$Var(X_t) = Var(Y + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t)$$

Since Y and ε_i are independent:

$$Var(X_t) = Var(Y) + Var(\varepsilon_1) + Var(\varepsilon_2) + \ldots + Var(\varepsilon_t)$$
$$= \sigma_Y^2 + \sigma^2 + \sigma^2 + \ldots + \sigma^2$$

(t terms)

$$= \sigma_Y^2 + t\sigma^2$$

This shows that the variance of X_t increases with time t, which violates the condition for weak stationarity that requires constant variance.

Therefore, the random walk model is nonstationary because its variance changes (specifically, grows) with time. This is a fundamental characteristic of random walks and explains why stock prices modeled in this way can drift arbitrarily far from their starting values over long time periods.

(2) Show that $Cov(X_t, X_s) = \sigma_Y^2 + \min(t, s)\sigma^2$.

To find $Cov(X_t, X_s)$, I'll use the expanded form of X_t and X_s in terms of the initial value Y and the white noise terms.

For any t:

$$X_t = Y + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_t$$

Similarly, for any s:

$$X_s = Y + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_s$$

Without loss of generality, let's assume $t \geq s$, so $\min(t, s) = s$.

The covariance is:

$$Cov(X_t, X_s) = Cov(Y + \varepsilon_1 + \ldots + \varepsilon_t, Y + \varepsilon_1 + \ldots + \varepsilon_s)$$

Using the bilinearity of covariance:

$$Cov(X_t, X_s) = Cov(Y, Y) + Cov(Y, \varepsilon_1 + \ldots + \varepsilon_s) + Cov(\varepsilon_1 + \ldots + \varepsilon_t, Y) + Cov(\varepsilon_1 +$$

Since Y is independent of all ε_i terms, $Cov(Y, \varepsilon_i) = 0$ and $Cov(\varepsilon_i, Y) = 0$.

Therefore:

$$Cov(X_t, X_s) = Var(Y) + Cov(\varepsilon_1 + \ldots + \varepsilon_t, \varepsilon_1 + \ldots + \varepsilon_s)$$

For the second term, we need to expand the covariance:

$$Cov(\varepsilon_1 + \ldots + \varepsilon_t, \varepsilon_1 + \ldots + \varepsilon_s) = \sum_{i=1}^{t} \sum_{j=1}^{s} Cov(\varepsilon_i, \varepsilon_j)$$

Since $\{\varepsilon_t\}$ is a white noise process, $Cov(\varepsilon_i, \varepsilon_j) = \sigma^2$ when $i = j$, and $Cov(\varepsilon_i, \varepsilon_j) = 0$ when $i \neq j$.

Thus, the double sum simplifies to the number of matching pairs:

$$Cov(\varepsilon_1 + \ldots + \varepsilon_t, \varepsilon_1 + \ldots + \varepsilon_s) = \sum_{i=1}^{s} \sigma^2 = s\sigma^2$$

(since $s \leq t$)

Therefore:

$$Cov(X_t, X_s) = Var(Y) + s\sigma^2 = \sigma_Y^2 + s\sigma^2 = \sigma_Y^2 + \min(t, s)\sigma^2$$

This completes the proof for $t \geq s$. Due to the symmetry of covariance, the result also holds when $s > t$, giving the general form:

$$Cov(X_t, X_s) = \sigma_Y^2 + \min(t, s)\sigma^2$$

(b) Let $\{\varepsilon_t\}$ be a white noise process of mean zero and variance σ^2, and $\{X_t\}$ is a moving average defined from it using the formula:

$$X_t = \frac{1}{2}\varepsilon_t + \frac{1}{3}\varepsilon_{t-1} + \frac{1}{6}\varepsilon_{t-2}.$$

Find the autocorrelation function of $\{X_t\}$.

To find the autocorrelation function of the moving average process $\{X_t\}$, I'll need to calculate the covariance between X_t and X_{t+h} for different lag values h, and then normalize by the variance of X_t.

First, let me identify what we're working with:

- $\{\varepsilon_t\}$ is a white noise process with mean 0 and variance σ^2
- X_t is defined as a weighted average: $X_t = \frac{1}{2}\varepsilon_t + \frac{1}{3}\varepsilon_{t-1} + \frac{1}{6}\varepsilon_{t-2}$

Step 1: Find the mean of X_t. Since ε_t has mean zero, and X_t is a linear combination of ε terms:

$$E[X_t] = \frac{1}{2}E[\varepsilon_t] + \frac{1}{3}E[\varepsilon_{t-1}] + \frac{1}{6}E[\varepsilon_{t-2}] = 0$$

Step 2: Find the variance of X_t. Since the ε terms are independent (white noise):

$$\mathrm{Var}(X_t) = \left(\frac{1}{2}\right)^2 \mathrm{Var}(\varepsilon_t) + \left(\frac{1}{3}\right)^2 \mathrm{Var}(\varepsilon_{t-1}) + \left(\frac{1}{6}\right)^2 \mathrm{Var}(\varepsilon_{t-2})$$

$$= \frac{1}{4}\sigma^2 + \frac{1}{9}\sigma^2 + \frac{1}{36}\sigma^2$$

$$= \frac{9}{36}\sigma^2 + \frac{4}{36}\sigma^2 + \frac{1}{36}\sigma^2$$

$$= \frac{14}{36}\sigma^2 = \frac{7}{18}\sigma^2$$

Step 3: Find the autocovariance function. For lag $h = 0$:

$$\gamma(0) = \mathrm{Var}(X_t) = \frac{7}{18}\sigma^2$$

For lag $h = 1$:

$$\gamma(1) = \mathrm{Cov}(X_t, X_{t+1})$$

$$= \mathrm{Cov}\left(\frac{1}{2}\varepsilon_t + \frac{1}{3}\varepsilon_{t-1} + \frac{1}{6}\varepsilon_{t-2}, \frac{1}{2}\varepsilon_{t+1} + \frac{1}{3}\varepsilon_t + \frac{1}{6}\varepsilon_{t-1}\right)$$

Since ε terms at different time points are uncorrelated (white noise property), only matching terms contribute:

$$\gamma(1) = \frac{1}{3} \cdot \frac{1}{2}\mathrm{Cov}(\varepsilon_t, \varepsilon_t) + \frac{1}{6} \cdot \frac{1}{3}\mathrm{Cov}(\varepsilon_{t-1}, \varepsilon_{t-1})$$

$$= \frac{1}{6}\sigma^2 + \frac{1}{18}\sigma^2$$

$$= \frac{3}{18}\sigma^2 + \frac{1}{18}\sigma^2 = \frac{4}{18}\sigma^2 = \frac{2}{9}\sigma^2$$

For lag $h = 2$:

$$\gamma(2) = \mathrm{Cov}(X_t, X_{t+2})$$

$$= \mathrm{Cov}\left(\frac{1}{2}\varepsilon_t + \frac{1}{3}\varepsilon_{t-1} + \frac{1}{6}\varepsilon_{t-2}, \frac{1}{2}\varepsilon_{t+2} + \frac{1}{3}\varepsilon_{t+1} + \frac{1}{6}\varepsilon_t\right)$$

Only the terms with ε_t match:

$$\gamma(2) = \frac{1}{2} \cdot \frac{1}{6}\mathrm{Cov}(\varepsilon_t, \varepsilon_t) = \frac{1}{12}\sigma^2$$

For lag $h = 3$ and beyond: There are no matching ε terms, so:

$$\gamma(h) = 0 \text{ for } h \geq 3$$

Step 4: Calculate the autocorrelation function $\rho(h) = \gamma(h)/\gamma(0)$.

$$\rho(0) = \frac{\gamma(0)}{\gamma(0)} = 1$$

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{\frac{2}{9}\sigma^2}{\frac{7}{18}\sigma^2} = \frac{2}{9} \times \frac{18}{7} = \frac{4}{7}$$

$$\rho(2) = \frac{\gamma(2)}{\gamma(0)} = \frac{\frac{1}{12}\sigma^2}{\frac{7}{18}\sigma^2} = \frac{1}{12} \times \frac{18}{7} = \frac{3}{14}$$

$$\rho(h) = 0 \text{ for } h \geq 3$$

Therefore, the autocorrelation function of $\{X_t\}$ is:

- $\rho(0) = 1$

- $\rho(1) = \frac{4}{7}$
- $\rho(2) = \frac{3}{14}$
- $\rho(h) = 0$ for $h \geq 3$

This is consistent with a moving average process of order 2 (MA(2)), which has non-zero autocorrelations only up to lag 2.

Question 3

(a) Suppose that $\{X_t\}$ follows the AR(1) process: $X_t = 0.3X_{t-1} + \varepsilon_t$, where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables with $\varepsilon_t \sim N(0, 6^2)$. If it is given that $X_t = 7$, find 95% confidence intervals for X_{t+1} and X_{t+2}.

To find the 95% confidence intervals for X_{t+1} and X_{t+2}, I need to determine their distributions given that $X_t = 7$.

Step 1: Find the distribution of X_{t+1} given $X_t = 7$.

Using the AR(1) model:

$$X_{t+1} = 0.3X_t + \varepsilon_{t+1}$$

Substituting $X_t = 7$:

$$X_{t+1} = 0.3 \times 7 + \varepsilon_{t+1} = 2.1 + \varepsilon_{t+1}$$

Since $\varepsilon_{t+1} \sim N(0, 6^2)$, we have:

$$X_{t+1} \sim N(2.1, 6^2)$$

For a 95% confidence interval, we need to find the values that capture the middle 95% of the distribution. Using the standard normal critical value $z_{0.025} = 1.96$:

$$X_{t+1} \in [2.1 - 1.96 \times 6, 2.1 + 1.96 \times 6]$$

$$X_{t+1} \in [2.1 - 11.76, 2.1 + 11.76]$$

$$X_{t+1} \in [-9.66, 13.86]$$

Therefore, the 95% confidence interval for X_{t+1} is $[-9.66, 13.86]$.

Step 2: Find the distribution of X_{t+2} given $X_t = 7$.

For X_{t+2}, I'll use the AR(1) model twice:

$$X_{t+2} = 0.3X_{t+1} + \varepsilon_{t+2}$$

Substituting the expression for X_{t+1}:

$$X_{t+2} = 0.3(0.3X_t + \varepsilon_{t+1}) + \varepsilon_{t+2}$$

$$X_{t+2} = 0.3 \times 0.3 \times 7 + 0.3\varepsilon_{t+1} + \varepsilon_{t+2}$$

$$X_{t+2} = 0.09 \times 7 + 0.3\varepsilon_{t+1} + \varepsilon_{t+2}$$

$$X_{t+2} = 0.63 + 0.3\varepsilon_{t+1} + \varepsilon_{t+2}$$

Since ε_{t+1} and ε_{t+2} are independent and normally distributed, their linear combination is also normally distributed with:

Mean = $0.63 + 0.3 \times E[\varepsilon_{t+1}] + E[\varepsilon_{t+2}] = 0.63 + 0 + 0 = 0.63$

Variance =
$(0.3)^2 \times Var(\varepsilon_{t+1}) + Var(\varepsilon_{t+2}) = (0.3)^2 \times 6^2 + 6^2 = 0.09 \times 36 + 36 = 3.24 + 36 =$

Therefore:

$$X_{t+2} \sim N(0.63, 39.24)$$

Standard deviation of $X_{t+2} = \sqrt{39.24} \approx 6.26$

The 95% confidence interval is:

$$X_{t+2} \in [0.63 - 1.96 \times 6.26, 0.63 + 1.96 \times 6.26]$$

$$X_{t+2} \in [0.63 - 12.27, 0.63 + 12.27]$$

$$X_{t+2} \in [-11.64, 12.90]$$

Therefore, the 95% confidence interval for X_{t+2} is $[-11.64, 12.90]$.

Summary

- 95% confidence interval for X_{t+1}: $[-9.66, 13.86]$
- 95% confidence interval for X_{t+2}: $[-11.64, 12.90]$

Note that the confidence interval for X_{t+2} is wider than for X_{t+1}, reflecting the increased uncertainty as we forecast further into the future.

(b)The values of sample autocorrelation function at lags 1-8 for a time series of 225 observations are shown in the following table:

k	1	2	3	4	5	6	7	8
$\hat{\rho}_k$	-0.35	-0.03	0.14	0.04	0.03	0.02	0.01	-0.01

Identify a moving average model for this dataset, and give your reasoning in detail.

To identify an appropriate moving average model, I'll analyze the sample autocorrelation function (ACF) pattern.

For a moving average process of order q (MA(q)), the theoretical autocorrelation function $\rho_k = 0$ for all $k > q$.

With 225 observations, the standard error of the sample ACF is approximately $1/\sqrt{225} \approx 0.067$. Using a 95% confidence interval, values exceeding $\pm 1.96 \times 0.067 \approx \pm 0.131$ are statistically significant.

Examining the sample ACF values:

- $\hat{\rho}_1 = -0.35$ is significant (negative)
- $\hat{\rho}_2 = -0.03$ is not significant
- $\hat{\rho}_3 = 0.14$ is just significant (positive)
- $\hat{\rho}_4$ through $\hat{\rho}_8$ are all not significant

This pattern suggests an MA(3) model since:

1. There are significant autocorrelations up to lag 3
2. The autocorrelations at lags 4 and beyond are not statistically significant

A general MA(3) model would be:

$$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3}$$

The unusual feature is that $\hat{\rho}_2$ is close to zero while $\hat{\rho}_1$ and $\hat{\rho}_3$ are significant. This suggests the parameters would need to satisfy the condition $\theta_2 + \theta_1\theta_3 \approx 0$ for the theoretical lag 2 autocorrelation to be near zero.

The negative value at lag 1 and positive value at lag 3 suggest that θ_1 would be negative and θ_3 would be positive in the fitted model.

Question 4

(a) Explain why "volatility clustering" can be seen in this model?

In the ARCH(1) model presented:

$$X_t = \sigma_t\varepsilon_t$$

$$\sigma_t^2 = 1 + 0.2X_{t-1}^2$$

Where $\{\varepsilon_t\}$ is an i.i.d sequence of standard normal random variables ($\varepsilon_t \sim N(0,1)$), and ε_t is independent of $\{X_s, s < t\}$.

Volatility clustering refers to the phenomenon where periods of high volatility tend to be followed by high volatility, and periods of low volatility tend to be followed by low volatility.

In this ARCH(1) model, volatility clustering is evident because the conditional variance σ_t^2 depends directly on the previous squared return X_{t-1}^2. Specifically:

1. If $|X_{t-1}|$ is large (a large positive or negative return), then X_{t-1}^2 is large, making σ_t^2 large, which increases the probability of another large return X_t.

2. If $|X_{t-1}|$ is small (a return close to zero), then X_{t-1}^2 is small, making σ_t^2 closer to 1, which increases the probability of another small return X_t.

This recursive relationship in the conditional variance equation creates persistence in volatility, which is exactly what volatility clustering describes.

(b) Derive the mean, variance, skewness and kurtosis of X_t Mean of X_t:

$$E[X_t] = E[\sigma_t \varepsilon_t] = E[\sigma_t] \cdot E[\varepsilon_t] = E[\sigma_t] \cdot 0 = 0$$

The mean is zero because ε_t has zero mean and is independent of σ_t.

Variance of X_t:

$$Var(X_t) = E[X_t^2] = E[\sigma_t^2 \varepsilon_t^2] = E[\sigma_t^2] \cdot E[\varepsilon_t^2] = E[\sigma_t^2] \cdot 1 = E[\sigma_t^2]$$

Since $\sigma_t^2 = 1 + 0.2X_{t-1}^2$, we have:

$$Var(X_t) = E[1 + 0.2X_{t-1}^2] = 1 + 0.2E[X_{t-1}^2]$$

For a stationary process, $E[X_t^2] = E[X_{t-1}^2] = Var(X_t)$, so:

$$Var(X_t) = 1 + 0.2Var(X_t)$$

$$0.8Var(X_t) = 1$$

$$Var(X_t) = 1.25$$

Skewness of X_t:

$$Skewness = \frac{E[X_t^3]}{(Var(X_t))^{3/2}}$$

$$E[X_t^3] = E[\sigma_t^3 \varepsilon_t^3] = E[\sigma_t^3] \cdot E[\varepsilon_t^3] = E[\sigma_t^3] \cdot 0 = 0$$

Therefore, $Skewness = 0$. The distribution is symmetric.

Kurtosis of X_t:

$$Kurtosis = \frac{E[X_t^4]}{(Var(X_t))^2} = \frac{E[X_t^4]}{(1.25)^2} = \frac{E[X_t^4]}{1.5625}$$

$$E[X_t^4] = E[\sigma_t^4 \varepsilon_t^4] = E[\sigma_t^4] \cdot E[\varepsilon_t^4] = E[\sigma_t^4] \cdot 3$$

To find $E[\sigma_t^4]$:

$$E[\sigma_t^4] = E[(1 + 0.2X_{t-1}^2)^2] = E[1 + 0.4X_{t-1}^2 + 0.04X_{t-1}^4]$$

$$= 1 + 0.4E[X_{t-1}^2] + 0.04E[X_{t-1}^4]$$

$$= 1 + 0.4 \cdot 1.25 + 0.04E[X_{t-1}^4]$$

$$= 1 + 0.5 + 0.04E[X_{t-1}^4]$$

$$= 1.5 + 0.04E[X_{t-1}^4]$$

For a stationary process, $E[X_t^4] = E[X_{t-1}^4]$, so:

$$E[X_t^4] = 3E[\sigma_t^4] = 3(1.5 + 0.04E[X_t^4])$$

$$E[X_t^4] = 4.5 + 0.12E[X_t^4]$$

$$0.88E[X_t^4] = 4.5$$

$$E[X_t^4] = 5.11$$

Therefore, $Kurtosis = \frac{5.11}{1.5625} = 3.27$

This is greater than 3 (the kurtosis of a normal distribution), indicating that the distribution has heavier tails than a normal distribution.

(c) Let $\eta_t = X_t^2 - \sigma_t^2$. Prove that $\{\eta_t\}$ has no autocorrelation.

To prove that $\{\eta_t\}$ has no autocorrelation, I need to show that $Cov(\eta_t, \eta_{t-k}) = 0$ for all $k > 0$.

First, let's express η_t in terms of the model:

$$\eta_t = X_t^2 - \sigma_t^2 = \sigma_t^2\varepsilon_t^2 - \sigma_t^2 = \sigma_t^2(\varepsilon_t^2 - 1)$$

Now, let's compute $Cov(\eta_t, \eta_{t-k})$ for $k > 0$:

$$Cov(\eta_t, \eta_{t-k}) = E[\eta_t\eta_{t-k}] - E[\eta_t]E[\eta_{t-k}]$$

First, I'll find $E[\eta_t]$:

$$E[\eta_t] = E[\sigma_t^2(\varepsilon_t^2 - 1)] = E[\sigma_t^2]E[\varepsilon_t^2 - 1] = E[\sigma_t^2] \cdot 0 = 0$$

Since $E[\varepsilon_t^2] = 1$ for standard normal variables, we have $E[\varepsilon_t^2 - 1] = 0$.

Now, let's compute $E[\eta_t\eta_{t-k}]$:

$$E[\eta_t\eta_{t-k}] = E[\sigma_t^2(\varepsilon_t^2 - 1) \cdot \sigma_{t-k}^2(\varepsilon_{t-k}^2 - 1)]$$

Since ε_t is independent of past values, $\varepsilon_t^2 - 1$ is independent of σ_t^2, σ_{t-k}^2, and $\varepsilon_{t-k}^2 - 1$ for $k > 0$.

Therefore:

$$E[\eta_t\eta_{t-k}] = E[\sigma_t^2\sigma_{t-k}^2]E[\varepsilon_t^2 - 1]E[\varepsilon_{t-k}^2 - 1] = E[\sigma_t^2\sigma_{t-k}^2] \cdot 0 \cdot 0 = 0$$

Thus, $Cov(\eta_t, \eta_{t-k}) = 0 - 0 = 0$ for all $k > 0$, which proves that $\{\eta_t\}$ has no autocorrelation.

This result is important because η_t represents the innovation or shock in the squared returns, and the lack of autocorrelation confirms that the ARCH(1) model has correctly captured the conditional heteroskedasticity structure in the data.