QUESTION 1

A time series is integrated of order d if $(1-L)^d X_t$ is a stationary process where L is the lag operator and $1-L$ is the first difference, i. e.

$$(1-L)X_t = X_t - X_{t-1} = \Delta X_t.$$

In other words, a process is integrated of order d if taking repeated differences d times yields a stationary process.

A process X_t is a autoregressive integrated moving average process *ARIMA (p,d,q)* if $Z_t \equiv \nabla^d X_t \equiv (1-L)^d X_t$ is *ARMA (p,q)*-process. So *ARIMA (p,d,q)* model is the model of the form

$$\nabla^d X_t = \phi_1 \nabla^d X_{t-1} + \ldots + \phi_p \nabla^d X_{t-p} + e_t + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}$$

where e_t is a white noise; ϕ_k and θ_j are real numbers (coefficients of autoregression and moving average respectively).

1. $Y_t = 2Y_{t-1} - Y_{t-2} + e_t$. This is *ARIMA (2,0,0)* model, i. e. *AR (2)*, where $\phi_1 = 2$, $\phi_2 = -1$.
2. $Y_t = Y_{t-1} - 0.25Y_{t-2} + e_t - 0.1e_{t-1}$. This is *ARIMA (2,0,1)* model, i. e. *ARMA (2,1)*, where $\phi_1 = 1$, $\phi_2 = -0.25$, $\theta_1 = -0.1$.
3. $Y_t = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$. This is *ARIMA (2,0,2)* model, i. e. *ARMA (2,2)*, where $\phi_1 = 0.5$, $\phi_2 = -0.5$, $\theta_1 = -0.5$, $\theta_2 = 0.25$.

QUESTION 2

We will use the definition of weak stationarity.

Random process is said to be weak stationary if:

- its expectation is constant: $E[X_t] = \mu_t = \mu = const.$
- autocovariance function does not depend on time: $\gamma(t_1, t_1 + \tau) = \gamma(t_2, t_2 + \tau) = \gamma(\tau)$.

If $\tau = 0$ $\gamma(0) = \sigma^2$ and we get that dispersion is also constant.

1. The graph shows gas consumption in UK from 1960Q1 to 1986Q4 (in millions of therms). The graph has rising trend. So the first condition of the definition above is violated (the expectation is not constant). The graph has seasonal fluctuations. So the second condition of the definition is violated too. The variance increases with time.
   The time series is not stationary.
2. We have a Q — Q plot. The x values are theoretical quantiles of a standard normal distribution and the y values are sample quantiles (the sample is the values of gas consumption). The points fall along a line in the middle of the graph, but curve off in the extremities. This behaviour means the data have more extreme values than would be expected if they truly came from a standard normal distribution. So the data do not have a standard normal distribution.
3. To remove the trend we can try differencing $\nabla X_t = X_t - X_{t-1}$. To remove the seasonal fluctuations we can apply seasonal defferencing $\nabla^4 X_t = X_t - X_{t-4}$ to the received series (we have the values for each quarter). We can use *log-return transformation* for stabilizing the variance. If it does not help we will apply *Box — Cox transformation.*

QUESTION 3

This is *ARIMA (2,0,1)* model, $\phi_1 = 1$, $\phi_2 = -0.49$, $\theta_1 = -0.56$. We will simulate in *R* with the help of *arima.sim()*. We will put $n = 100$ and we will get $E(X_t) = 0.03$, $Var(X_t) = 1.4$.

QUESTION 4

The original series is not stationary. We see it from its graph (the expectation is not constant). After first differencing we get the stationary time series. The graph of the received series looks like a white noise. The graphs of ACF and $PACF$ are typical for stationary time series.

From the graph of $PACF$ we see that $p = 1$ and from the graph of ACF we see that $q = 1$. As we look at the time series after first differencing $d = 1$. So this is *ARIMA (4,1,1)* model:

$$\Delta Y_t = \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \phi_3 \Delta Y_{t-3} + \phi_4 \Delta Y_{t-4} + e_t + \theta_1 e_{t-1}.$$