Question 1 [20 points]:

1. The conditional variance a time $T+1$ is

$$\sigma^2_{T+1} = 0.1 + 0.5y^2_T + 0.3y^2_{T-1} = 1.8$$

2. The conditional distribution of y_{T+1} is $y_{T+1}|Y^T \sim N(0, \sigma^2_{T+1})$. Therefore,

$$P(y_{T+1} < -1.0|Y^T) = P\left(\frac{y_{T+1}}{\sigma_{T+1}} < \frac{-1.0}{\sigma_{T+1}}\right) = \Phi\left(\frac{-1.0}{\sigma_{T+1}}\right) = \Phi(-0.75).$$

Note: using -0.1 instead of -1.0 was also considered correct as there was a typo in text.

3. When $y_T = -1.5$, we obtain a larger value of σ^2_{T+1} since y^2_T is larger. Therefore, the probability will be larger since the variance of the normal will be larger and hence the probability of tail events will be larger.

Question 2 [10 points]:
The index of `sig2[t]` may be changed to `t+1`, although this is not necessarily an error as long as the rest of the code (which is not present in the question) takes this into account. The code for an ARCH(3) would be as follows

```
omega <- exp(par[1])
alpha1 <- exp(par[2])/(1+exp(par[2]))
alpha2 <- exp(par[3])/(1+exp(par[3]))
alpha3 <- exp(par[4])/(1+exp(par[4]))

sig2 <- rep(var(x),n)

for(t in 3:n){

  sig2[t] <- omega + alpha1 * x[t]^2 + alpha2 * x[t-1]^2 + alpha3 * x[t-2]^2

}
```

Question 3 [25 points]:

1. The conditional variances and covariance at time $T+1$ are $\sigma^2_{1,T+1} = 3$, $\sigma^2_{2,T+1} = 1.5$, and $\sigma_{12,T+1} = -0.5$. Therefore, the conditional variance of the portfolio is

$$\sigma^2_{p,T+1} = 0.4^2 \times 3 + 0.6^2 \times 2.5 + 2 \times 0.4 \times 0.6 \times (-0.5) = 1.14.$$

Therefore, the VaR is

$$1\%\text{-VaR}_{T+1} = z_{0.01} \times \sigma_{p,T+1} = -2.33 \times \sqrt{1.14} = -2.49.$$

2. The unconditional variance of the assets is

$$\Sigma = W(1-\alpha)^{-1}\begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix} \times 2 = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}.$$

Therefore, the unconditional variance of the portfolio is

$$\sigma^2_p = 0.4^2 \times 2 + 0.6^2 \times 4 + 2 \times 0.4 \times 0.6 \times 1 = 2.24.$$

3. It is true that the number of parameters in the optimization of the likelihood is two. However, the number of parameters in the model will still be $n \times (n+1)/2 + 2$. Covariance targeting does not improve estimation accuracy, standard ML is asymptotically efficient. The model with more parameters than observations cannot be estimated even with covariance targeting. The problem is that the estimated unconditional covariance matrix will be singular (not positive definite). Assume that Y is an $T \times n$ matrix that contains the log-returns. Then, the estimated covariance matrix $\hat{\Sigma} = Y^\top Y$ will have rank $T < n$, and hence it will be singular.

Question 4 [18 points]:

1. The statement is correct. In general AIC and BIC do not give information on correct specification or misspecification, but in this case they do. If the GARCH(2,2) has lower AIC and BIC compared to GARCH(1,1), it means that the likelihood of the GARCH(2,2) is significantly larger than the one of the GARCH(1,1) to compensate the penalty for additional parameters. This indicates misspecification of the GARCH(1,1) since if the GARCH(1,1) is correctly specified it should have the same likelihood of GARCH(2,2) in the limit. Therefore, the larger likelihood of GARCH(2,2) in small sample will tend not to compensate the penalty for additional parameters.

2. The statement is not correct. The test should be based on the residuals and not on the log-returns directly. The reason is that conditional heteroschestaticity creates heavy tails in the unconditional distribution of the log-returns. Therefore, when testing normality on the log-returns, we will tend to reject the null even when the error is normal.

Question 5 [17 points]:

1. First, we notice that the unconditional mean of y_t is zero

$$\mathbb{E}(y_t) = \mathbb{E}(\beta_t x_t) + \mathbb{E}(\epsilon_t) = \mathbb{E}(\beta_t)\mathbb{E}(x_t) = 0,$$

where the second equality follows since x_t and η_t are independent sequences. Next, we obtain that

$$\mathbb{C}ov(y_t, y_{t-k}) = \mathbb{E}(y_t y_{t-l}) = \mathbb{E}(\beta_t x_t \beta_{t-k} x_{t-k}) = \mathbb{E}(\beta_t \beta_{t-k})\mathbb{E}(x_t)\mathbb{E}(x_{t-k}) = 0,$$

where the second equality follows since x_t is iid and independent of η_t, and the third equality follows sine x_t has mean zero. Therefore, the autocovariance function is zero at any lag. Finally, we notice that if ϕ_1 and ϕ_2 are different from zero, then y_t is not independent of y_{t-1} and y_{t-2} since f_t is an MA(2) and hence y_t,y_{t-1} and y_{t-2} all depend on η_{t-2}. Instead, y_t is independent of y_{t-k} for $k > 2$ since y_{t-k} will not depend on η_t, η_{t-1}, and η_{t-2}.

2. First, we obtain the conditional mean

$$\mathbb{E}(y_t|x_t) = \mathbb{E}(\beta_t x_t|x_t) + \mathbb{E}(\epsilon_t|x_t) = \mathbb{E}(\beta_t)x_t.$$

We notice that $f_t \sim N(0, \sigma_f^2)$, where $\sigma_f^2 = 1 + \phi_1^2 + \phi_2^2$. Therefore,

$$\mathbb{E}(y_t|x_t) = \exp(\sigma_f^2/2)x_t.$$

Next, we obtain the conditional second moment

$$\mathbb{E}(y_t^2|x_t) = \mathbb{E}(\beta_t^2 x_t^2|x_t) + \mathbb{E}(\epsilon_t^2|x_t) + 2\mathbb{E}(\beta_t x_t \epsilon_t|x_t) = \mathbb{E}(\beta_t^2)x_t^2 + 1.$$

Next, we notice that $\beta_t^2 = \exp(2f_t)$ and $2f_t \sim N(0, 4\sigma_f^2)$. Therefore,

$$\mathbb{E}(y_t^2|x_t) = \exp(2\sigma_f^2)x_t^2 + 1.$$

Therefore the conditional variance is

$$\mathbb{V}ar(y_t|x_t) = \mathbb{E}(y_t^2|x_t) - \mathbb{E}(y_t|x_t)^2 = x_t^2 \exp(\sigma_f^2)\left(\exp(\sigma_f^2) - 1\right) + 1.$$

Question 6 [10 points]:
The statement is not correct. There are two problems. The first is that it is not true that we necessarily need to use 3 auxiliary statistics. The number of auxiliary statistics can be larger equal than the number of parameters. Furthermore, using the set of auxiliary statistics proposed we will not give a consistent estimator since the unconditional mean of an SV model is zero for any parameter value and therefore the sample mean will not identify any parameter. Therefore, in fact, we will be using only 2 auxiliary statistics. *Note: some of you may have interpreted the question as the sample mean of y_t^2, which would identify the parameters. That answer was considered correct as well.*