Question 1(a)

Since r_t, r_{t-1}, r_{t-2} are i.i.d and follow normal distribution.

$r_t + r_{t-1} + r_{t-2} = r_t(3)$ also follows normal distribution. Also:

$$r_t(3) \sim \mathcal{N}(\mu_{r_t} + \mu_{r_{t-1}} + \mu_{r_{t-2}},\, \sigma^2_{\mu} r_t + \sigma^2_{\mu} r_{t-1} + \sigma^2_{\mu} r_{t-2})$$

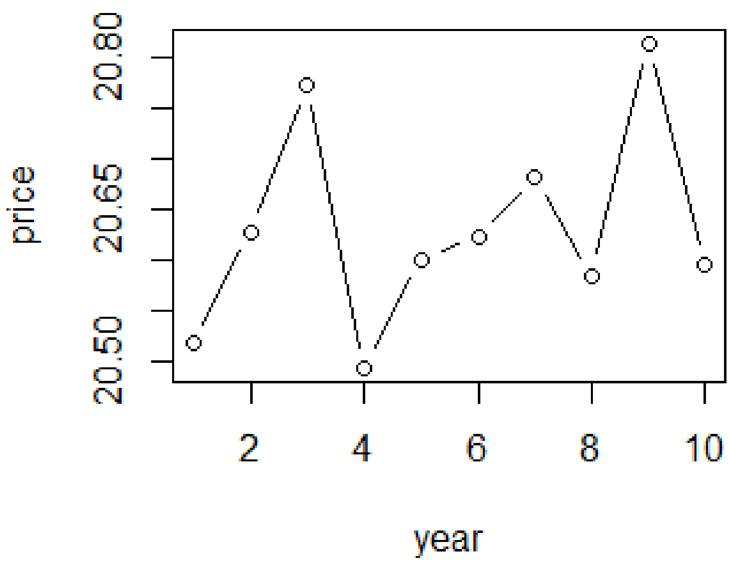
So $r_t(3) \sim \mathcal{N}(3\mu, 3\sigma^2)$

Question 1(b)

Since r_t, r_{t-1}, r_{t-2} are i.i.d and follow $\mathcal{N}(\mu, \sigma^2)$. Co-variance of any of the 2 different terms is zero (*). Also, variance of any term is σ^2. Then:

$$
\begin{aligned}
&COV[r_t(k), r_t(k+l)] \\
&= COV[r_t + r_{t-1} + ... + r_{t-k}, [r_t + r_{t-1} + ... + r_{t-k-l}] \\
&= Var(r_t) + ... + Var(r_{t-k}) \qquad (By\ *) \\
&= k\sigma^2
\end{aligned}
$$

Question 2




```
#2(a)
p0 <- 20
mu <- 0.03
sd = 0.005
sum_ <- rnorm(10, mu, sd)
price <- p0*exp(sum_)
time <- 1:10
plot(time, price, type = "b", xlab = "year", ylab = "price")

#2(b)
seq <- rep()
for(i in 1:2000){
  sum2 <- rnorm(1,mu,sd)
  seq[i]=p0*exp(sum2)
  #print(seq[i])
}

mean_seq <- mean(seq)
print(mean_seq)

#2(c)
mean_p10 <- p0*exp(mu+sd^2/2)
print(mean_p10)
```

From the r code above, expected mean of P_{10} is 20.61354. The simulated mean is 20.60935. They are almost the same.

Question 3

Convex pattern indicates right skewness of the data. Concave pattern indicates left skewness. Convex-concave pattern indicates light tails compare with normal distribution. Concave-convex pattern indicates heavy tails compare with normal distribution.

Question 4

(a)

$$
\begin{aligned}
X_2 &= X_0 * exp(r_1 + r_2) \\
&= X_0 * (x_2/x_0)
\end{aligned}
$$

where $r_1 + r_2 = log(x_2/x_0) \sim \mathcal{N}(2\mu, 2\sigma^2) since \mathrm{r}_1, r_2$ follows lognormal and are i.i.d. Then:

$$
\begin{aligned}
P(x_2 > 1.5x_0) &= P((x_2/x_0) > 1.5) \\
&= P(r_1 + r_2 > log(1.5)) \qquad (Log\ Both\ Sides) \\
&= 1 - \Phi(\frac{log(1.5) - 2\mu}{\sqrt{2\sigma}})
\end{aligned}
$$

(b)

Let q be the value for the 0.8 quantile of $X-k$ for all k. I want to find a formula for q. Then:

$$\begin{aligned} P(X_k < q) &= 0.8 \\ P(log(X_k) < log(q)) &= 0.8 \qquad (Log\ Both\ Sides) \\ log(q) - log(X_0) - k\mu &= 1.28\sqrt{k\sigma} \\ log(q) &= 1.28\sqrt{k\sigma} + log(X_0) + k\mu \end{aligned}$$

So the formula for q is $X_0 * exp(k\mu + 1.28\sqrt{k\sigma})$.

(c)

$X_k = X_0 * exp(r_1 + ... + r_k)$ (*). Now consider $log(X_K)$. It's in $\mathcal{N}(log(X_0) + k\mu, k\sigma^2)$ Then:

$$p(X_K) = \frac{1}{\sqrt{2\pi k\sigma X_k}} exp(\frac{[log(X_0) + k\mu]^2 - log(X_k)}{2k\sigma^2})$$

$$\begin{aligned} E(X_k^2) &= \int_{-\infty}^{\infty} X_k^2 * p(X_k) dX_k \\ &= X_0^2 * exp(2k\mu + 2k\sigma^2) * \int_{-\infty}^{\infty} p(r_1 + ... + r_k) d(r_1 + ... + r_k) \qquad (By\ *) \\ &= X_0^2 * exp(2k\mu + 2k\sigma^2) * 1 \end{aligned}$$

(d)

Similarly,

$$
\begin{aligned}
E(X_k) &= \int_{-\infty}^{\infty} X_k * p(X_k) dX_k \\
&= X_0^2 * exp(k\mu - \frac{k\sigma^2}{2}) \\
Var(X_k) &= E[X_k^2](E[X_k])^2 \\
&= X_0^2 * exp(2k\mu + 2k\sigma^2) - [X_0^2 * exp(k\mu - \frac{k\sigma^2}{2})]^2
\end{aligned}
$$

Question 5

The likelihood function (log) is:

$$L(\mu, \sigma^2; Y_1, ..., Y_n) = -n/2ln(\sigma^2) - n/2ln(2\pi) - 1/2\sigma^2 \sum_{i=1}^{n} (Y_i - \mu)^2$$

So the objective function is:

$$\max_{\sigma^2} L(\mu, \sigma^2; Y_1, ..., Y_n)$$

Subject to:

$$\frac{\partial}{\partial \sigma^2} L(\mu, \sigma^2; Y_1, ..., Y_n) = 0$$

Then:

$$\frac{\partial}{\partial \sigma^2} L(\mu, \sigma^2; Y_1, ..., Y_n)$$
$$= \frac{\partial}{\partial \sigma^2} - n/2ln(\sigma^2) - n/2ln(2\pi) - 1/2\sigma^2 \sum_{i=1}^{n} (Y_i - \mu)^2$$
$$= (1/2)[\sum_{i=1}^{n} (Y_i - \mu)^2 (1/\sigma^4) - n/2\sigma^2$$
$$= (1/\sigma^2)[1/\sigma^2 \sum_{i=1}^{n} (Y_i - \mu)^2] - n]$$

Assuming σ^2 is not zero, $1/\sigma^2[\sum_{i=1}^{n}(Y_i - \mu)^2] - n = 0$ Then:

$$1/\hat{\sigma}^2[\sum_{i=1}^{n} (Y_i - \mu)^2] = n$$

$$\hat{\sigma}^2 = n^{-1} \sum_{i=1}^{n} (Y_i - \mu)^2$$

Question 6

(a)

$$Z = pX_1 + (1 - p)X_2$$

$$Mean(Z) = 0$$

$$Var(Z) = p * \sigma_1^2 + (1 - p) * \sigma_2^2$$

(b)

Intuitively, mixture models have heavier tails than normal distribution with the same mean and variance. Since variables in the mixture model are more likely to be in the outlier range. I want to calculate the excess kurtosis to confirm this.

$$excess\_kurtosis = 3[\frac{(p * \sigma_1^4 + (1-p) * \sigma_2^4)}{(p * \sigma_1^2 + (1-p) * \sigma_2^2)^2}] - 3$$

From the above result, excess kurtosis equals zero only if p = 0 or 1, meaning the mixture model does not have heavier tails compare with normal distribution with the same mean and variance. Otherwise excess kurtosis is greater than zero, meaning mixture model has heavier tails.