# Assignment 1

Yongseok Kim

February 12, 2020

1. (a) Table 1 reports summary statistics of the data.

   (b) Let the excess return process $X_t =_{i.i.d} \mathbb{N}(\mu, \sigma^2)$ for $t = 1, \cdots, T$. Then, we have the log likelihood function of $X$

   $$l(\mu, \sigma) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}(X - \mu\iota)'(X - \mu\iota)$$

   Then, the MLE estimators for $(\mu, \sigma^2)$ yield

   $$\hat{\mu} = \frac{1}{T}\sum_t x_t$$

   and

   $$\hat{\sigma}^2 = \frac{1}{T}\sum_t (X_t - \hat{\mu})^2$$

   Note that

   $$\begin{aligned}\hat{\mu} &= \frac{1}{T}\iota' X \\ &=_d \frac{1}{T}\iota'\mathbb{N}(\mu\iota, \sigma^2 I) \\ &= \mathbb{N}(\mu, \sigma^2/T)\end{aligned}$$

   and

   $$\mathbb{E}(X_i - \mu)^2 = \sigma^2$$

Table 1: Summary Statistics
In this table, I report number of observations (N), sample mean (Mean), standard deviation (St. Dev.), minimum (Min), the first quartile (Pctl(25)), median (Median), the third quartile (Pctl(75)), and maximum (Max) of each variables. Mkt-RF denotes the excess market returns, SMB and HML denote size and book-to-market factor from Fama and French (1993), and RF denotes monthly risk-free rates.

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Median | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|---|
| Mkt-RF | 1,108 | 0.659 | 5.330 | −29.130 | −1.970 | 1.015 | 3.615 | 38.850 |
| SMB | 1,108 | 0.207 | 3.192 | −16.870 | −1.562 | 0.070 | 1.730 | 36.700 |
| HML | 1,108 | 0.369 | 3.484 | −13.280 | −1.320 | 0.140 | 1.740 | 35.460 |
| RF | 1,108 | 0.274 | 0.253 | −0.060 | 0.030 | 0.230 | 0.430 | 1.350 |

$$\mathrm{var}(X_i - \mu)^2 = \mathbb{E}(X_i - \mu)^4 - \sigma^4 = 2\sigma^4$$

imply

$$\begin{aligned}
\sqrt{T}(\hat{\sigma}^2 - \sigma^2) &= \sqrt{T}\big(\frac{1}{T}\sum_t (X_t - \mu)^2 - \sigma^2 - (\mu - \hat{\mu})^2\big) \\
&= \frac{1}{\sqrt{T}}\sum_t \big((X_t - \mu)^2 - \sigma^2\big) - \frac{1}{\sqrt{T}}\big(\sqrt{T}(\mu - \hat{\mu})\big)^2 \\
&\to_d \mathbb{N}(0, 2\sigma^4)
\end{aligned}$$

$$\Rightarrow \hat{\sigma}^2 \approx_d \mathbb{N}(\sigma^2, \frac{2\sigma^4}{T})$$

for large $T$. Since

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{T}\sum_t (X_t - \hat{\mu})^2 \\
&= \frac{1}{T}\sum_t (X_t - \mu)^2 - (\mu - \hat{\mu})^2 \\
&\to_p \mathbb{E}(X_t - \mu)^2 \\
&= \sigma^2
\end{aligned}$$

by the WLLN, we may approximate

$$\mathrm{var}(\hat{\mu}) \approx \hat{\sigma}^2/T$$

and

$$\mathrm{var}(\hat{\sigma}^2) \approx 2(\hat{\sigma}^2)^2/T$$

Finally, from the data, we obtain

$$\hat{\mu} = 0.659$$

$$\hat{\sigma}^2 = 28.381$$

$$\mathrm{var}(\hat{\mu}) \approx 0.026$$

$$\mathrm{se}(\hat{\mu}) \approx 0.160$$

$$\mathrm{var}(\hat{\sigma}^2) \approx 1.454$$

$$\mathrm{se}(\hat{\sigma}^2) \approx 1.206$$

(c) Assume $(X_t)$ is i.i.d. distributed with $\mathbb{E}X_t = \mu$ and $\mathrm{var}(X_t) = \sigma^2$. We have, by the WLLN,

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{T}\sum_t (X_t - \hat{\mu})^2 \\
&= \frac{1}{T}\sum_t (X_t - \mu)^2 - (\hat{\mu} - \mu)^2 \\
&\to_p \mathbb{E}(X_t - \mu)^2 \\
&= \sigma^2
\end{aligned}$$

and by the CLT

$$\sqrt{T}(\hat{\mu} - \mu) \approx_d \mathbb{N}(0, \sigma^2)$$
$$\Rightarrow \hat{\mu} \approx_d \mathbb{N}(\mu, \sigma^2/T)$$

Consider a continuous function

$$f_{\sigma^2}(x) = \frac{x}{\sqrt{\sigma^2}}$$

and first order derivative of $f$

$$df_{\sigma^2}(x) = 1/\sqrt{\sigma^2}$$

which is also continuous. Then, by the Delta method, we have

$$\begin{aligned}\frac{\hat{\mu}}{\sigma} - \frac{\mu}{\sigma} &\approx_d \frac{1}{\sqrt{\sigma^2}}\mathbb{N}(0, \sigma^2/T) \\ &= \mathbb{N}(0, \frac{1}{T})\end{aligned}$$

Finally, since $\hat{\sigma}^2 = \sigma^2 + o_p(1)$, we have

$$\frac{\hat{\mu}}{\hat{\sigma}} - \frac{\mu}{\sigma} \approx_d \mathbb{N}(0, \frac{1}{T})$$

Hence, we may approximate

$$\begin{aligned}\text{se}(\frac{\hat{\mu}}{\hat{\sigma}}) &\approx 1/T \\ &= 0.001\end{aligned}$$

The numerical Delta method is following. Let

$$df_{\sigma^2}(x) = \frac{f_{\sigma^2}(x+h) - f_{\sigma^2}(x-h)}{2h}$$

for sufficiently small $h$. Then, we may approximate

$$\begin{aligned}\text{se}(\frac{\hat{\mu}}{\hat{\sigma}}) &\approx \sqrt{\frac{\hat{\sigma}^2}{T} df_{\hat{\sigma}^2}(\hat{\mu})^2} \\ &= 0.001\end{aligned}$$

2. Let $X_n =_{i.i.d} \mathbb{N}(\mu, \sigma^2)$ where $\mu = 1$, $\sigma^2 = 4$ and $n = 100$. The sample log likelihood function for a univariate normal distribution with mean $\mu$ and standard deviation $\sigma$ is

$$\begin{aligned}l(x_i, \theta) &= \log p(x_i, \theta) \\ &= -\frac{1}{2}\log 2\pi - \frac{1}{2}\log \sigma^2 - \frac{1}{2\sigma^2}(x_i - \mu)^2\end{aligned}$$

where $p(X, \theta)$ is a density function of $X =_d \mathbb{N}(\mu, \sigma^2)$. The mean log likelihood is defined as

$$\frac{1}{n}\sum_{i=1}^{n} l(x_i, \theta).$$

For the numerical solution of the maximum likelihood estimate from $X_n$, see Figure 1.

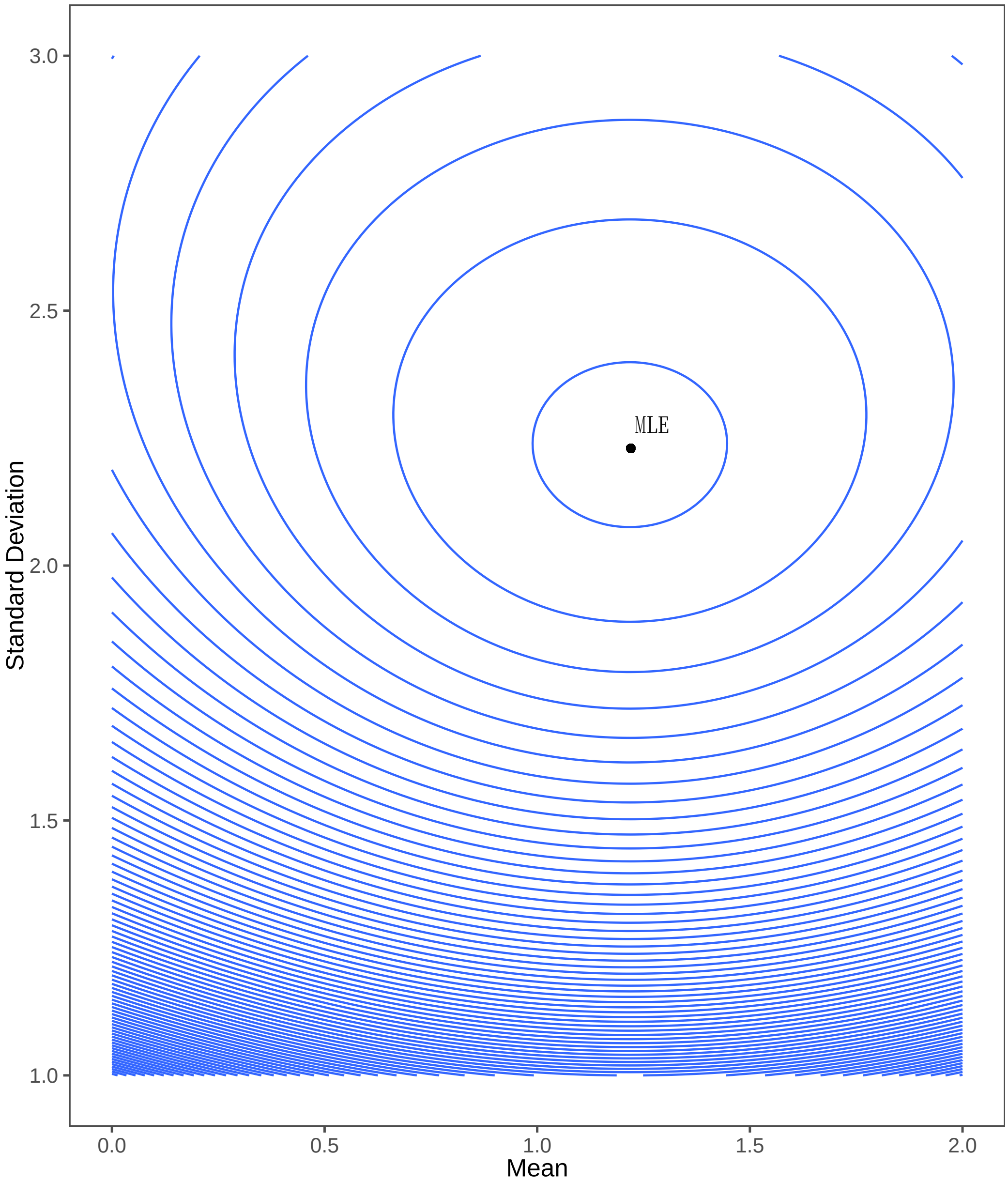


Figure 1: The contour of the mean log-likelihood. $\hat{\mu}_{ML} = 1.22$, $\hat{\sigma}_{ML} = 2.23$, and $l(x, \hat{\mu}_{ML}, \hat{\sigma}_{ML}) = -2.22$.

Table 2: Summary Statistics

| Statistic | N | Mean | St. Dev. | Min | Pctl(25) | Median | Pctl(75) | Max |
|---|---|---|---|---|---|---|---|---|
| Mkt-RF | 1,108 | 0.659 | 5.330 | −29.130 | −1.970 | 1.015 | 3.615 | 38.850 |
| SMB | 1,108 | 0.207 | 3.192 | −16.870 | −1.562 | 0.070 | 1.730 | 36.700 |
| HML | 1,108 | 0.369 | 3.484 | −13.280 | −1.320 | 0.140 | 1.740 | 35.460 |

Table 3: OLS estimation results of Eq (1)

| param | estimate | se | tstat | pvalue |
|---|---|---|---|---|
| alpha | 0.082 | 0.092 | 0.892 | 0.372 |
| beta | 0.191 | 0.017 | | |

3. (a) Table 2 show summary statistics of the excess market returns, SMB, and HML. Three portfolios yield positive average returns. Measuring performances by the Sharpe ratio, the excess returns do the best during the sample period, followed by HML.

   (b) For OLS estimation results of Equation (1), see Table 3. Table 4 reports OLS estimation results of Equation (2).

   (c) Table 5 and 6 report the OLS estimation results of Equation (1) and (2), respectively. Standard errors are smaller and $t$-statstics are greater than (b).

   (d) Constructing Wald statistics for the null hypotheses from Equation (3) and (4) is

$$W = \frac{(RSS_1 - RSS_0)/l}{RSS_0/(n-k)} =_d F(l, n-k)$$

   where $l$ is number of restrictions, $n$ is number of observations, and $k$ is number of regressors. For Equation (5), stacking Equation (1) and (2) yields

$$\begin{bmatrix} r_{SMB} \\ r_{HML} \end{bmatrix} = \alpha_{SMB} \begin{bmatrix} \iota \\ 0 \end{bmatrix} + \alpha_{HML} \begin{bmatrix} 0 \\ \iota \end{bmatrix} + \beta_{SMB} \begin{bmatrix} r_M \\ 0 \end{bmatrix} + \beta_{HML} \begin{bmatrix} 0 \\ r_M \end{bmatrix} + \begin{bmatrix} \epsilon_{SMB} \\ \epsilon_{HML} \end{bmatrix}$$

   Assume

$$\text{var}(\begin{bmatrix} \epsilon_{SMB} \\ \epsilon_{HML} \end{bmatrix} | X) = \sigma^2 I$$

   Then constructing Wald statistics for the null hypotheses from Equation (5) becomes clear.
   Table 7 report Wald statistics and $p$-values from $F$-distribution. For Equation (3)

Table 4: OLS estimation results of Eq (2)

| param | estimate | se | tstat | pvalue |
|---|---|---|---|---|
| alpha | 0.268 | 0.102 | 2.612 | 0.009 |
| beta | 0.154 | 0.019 | | |

Table 5: OLS estimation results of Eq (1) with White (1980) standard errors

| param | estimate | se | tstat | pvalue |
|---|---|---|---|---|
| alpha | 0.082 | 0.085 | 0.960 | 0.337 |
| beta | 0.191 | 0.033 | | |

Table 6: OLS estimation results of Eq (2) with White (1980) standard errors

| param | estimate | se | tstat | pvalue |
|---|---|---|---|---|
| alpha | 0.268 | 0.096 | 2.793 | 0.005 |
| beta | 0.154 | 0.054 | | |

and (4), $p$-values are different from (c) since Wald statistics do not correct conditional heteroskedasticity. However, Wald statistics can test a joint hypothesis like Equation (5), which cannot be done by $t$-statistics.

4. (a) Define

$$h^1(\theta, r_{M,t}) = r_{M,t} - \mu$$
$$h^2(\theta, r_{M,t}) = (r_{M,t} - \mu)^2 - \sigma^2.$$

Since the parameters are exactly identified, we set the sample moment conditions to zero

$$\tilde{h}_T^1(\theta) = \frac{1}{T}\sum r_{M,t} - \hat{\mu} = 0,$$
$$\tilde{h}_T^2(\theta) = \frac{1}{T}\sum (r_{M,t} - \hat{\mu})^2 - \hat{\sigma}^2 = 0.$$

Then, we obtain

$$\hat{\mu} = \frac{1}{T}\sum r_{M,t}$$
$$= 0.6590,$$

$$\hat{\sigma}^2 = \frac{1}{T}\sum (r_{M,t} - \hat{\mu})^2$$
$$= 28.3814$$

Table 7: Wald statistics and $p$-values

| eq | wald | pvalue |
|---|---|---|
| 3 | 0.795 | 0.373 |
| 4 | 6.810 | 0.009 |
| 5 | 4.139 | 0.016 |

Under the no autocorrelation assumption, we obtain

$$
\begin{aligned}
S_T &= \frac{1}{T}\sum_t h(\theta, r_{M,t})h(\theta, r_{M,t})' \\
&= \begin{bmatrix} \frac{1}{T}\sum_t (r_{M,t}-\hat{\mu})^2 & \frac{1}{T}\sum_t (r_{M,t}-\hat{\mu})\big((r_{M,t}-\hat{\mu})^2-\hat{\sigma}^2\big) \\ \frac{1}{T}\sum_t (r_{M,t}-\hat{\mu})\big((r_{M,t}-\hat{\mu})^2-\hat{\sigma}^2\big) & \frac{1}{T}\sum_t \big((r_{M,t}-\hat{\mu})^2-\hat{\sigma}^2\big)^2 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
D_T &= \begin{bmatrix} -1 & 0 \\ -\frac{2}{T}\sum_t (r_{M,t}-\hat{\mu}) & -1 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
V_T &= [D_T S_T^{-1} D_T']^{-1} \\
&= S_T
\end{aligned}
$$

$$
\Rightarrow \begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{bmatrix} \approx \mathbb{N}\big(\begin{bmatrix} \mu \\ \sigma^2 \end{bmatrix}, S_T/T\big)
$$

As a result, we obtain

$$
\begin{aligned}
\mathrm{se}(\hat{\mu}) &= \sqrt{\frac{1}{T}\sum_t (r_{M,t}-\hat{\mu})^2/T} \\
&= 0.160
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{se}(\hat{\sigma}^2) &= \sqrt{\frac{1}{T}\sum_t \big((r_{M,t}-\hat{\mu})^2-\hat{\sigma}^2\big)^2/T} \\
&= 2.681
\end{aligned}
$$

(b) Suppose

$$
r_{M,t} =_{i.i.d} \mathbb{N}(\mu, \sigma^2).
$$

Then

$$
\begin{aligned}
\hat{\mu} &= \frac{1}{T}\iota' r_M \\
&=_d \mathbb{N}(\mu, \sigma^2/T)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{T\hat{\sigma}^2}{\sigma^2} &= \frac{r_M'}{\sigma}(I-P_\iota)\frac{r_M}{\sigma} \\
&= \big(\frac{r_M-\mu}{\sigma}\big)'(I-P_\iota)\big(\frac{r_M-\mu}{\sigma}\big) \\
&=_d \chi^2_{T-1}
\end{aligned}
$$

$$
\Rightarrow \mathrm{var}(\hat{\sigma}^2) = \frac{2(T-1)}{T^2}\sigma^4
$$

Moreover, due to the normality, we have

$$\hat{\mu} = f(\iota' r_M)$$

and

$$\hat{\sigma}^2 = g\big((I - P_\iota) r_M\big)$$

are independent. Hence, we obtain

$$\mathrm{var}\Big( \begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \end{bmatrix} \Big) = \begin{bmatrix} \sigma^2/T & 0 \\ 0 & 2(T-1)\sigma^4/T^2 \end{bmatrix}$$

Since $\hat{\sigma}^2 \to_p \sigma^2$, we may approximate that

$$\begin{aligned} \mathrm{se}(\hat{\mu}) &\approx \sqrt{\hat{\sigma}^2/T} \\ &= 0.160 \end{aligned}$$

$$\begin{aligned} \mathrm{se}(\hat{\sigma}^2) &\approx \sqrt{2(T-1)(\hat{\sigma}^2)^2/T^2} \\ &= 1.205 \end{aligned}$$

(c) First, note that

$$\Gamma_{1,T} = \frac{1}{T} \sum_{t=2}^{T} h(\theta, r_{M,t}) h(\theta, r_{M,t-1})'$$

$$= \begin{bmatrix} \frac{1}{T}\sum_{t=2}^{T}(r_{M,t} - \hat{\mu})(r_{M,t-1} - \hat{\mu}) & \frac{1}{T}\sum_{t=2}^{T}(r_{M,t} - \hat{\mu})\big((r_{M,t-1} - \hat{\mu})^2 - \hat{\sigma}^2\big) \\ \frac{1}{T}\sum_{t=2}^{T}(r_{M,t-1} - \hat{\mu})\big((r_{M,t} - \hat{\mu})^2 - \hat{\sigma}^2\big) & \frac{1}{T}\sum_{t=2}^{T}\big((r_{M,t} - \hat{\mu})^2 - \hat{\sigma}^2\big)\big((r_{M,t-1} - \hat{\mu})^2 - \hat{\sigma}^2\big) \end{bmatrix}$$

Then

$$\tilde{S}_T = S_T + (1 - 1/2)\big(\Gamma_{1,T} + \Gamma_{1,T}'\big)$$

and

$$\tilde{V}_T = \tilde{S}_T$$

where $S_T$ is from part (a). Then, we obtain

$$\mathrm{se}(\hat{\mu}) = 0.169$$

$$\mathrm{se}(\hat{\sigma}^2) = 3.049$$

which is greater than standard errors calculated in part (a).

5. First, note that a stochastic process $(x_t)$ has

$$\mathbb{E}x_t = \theta \frac{1}{1-\rho}$$

$$\mathrm{var}(x_t) = \sigma^2 \frac{1}{1-\rho^2}$$

$$\gamma(k) = \rho^{|k|} \sigma^2 \frac{1}{1-\sigma^2}$$

suggesting it is covariance stationary. Define

$$y = \begin{bmatrix} x_1 \\ \vdots \\ x_T \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & x_0 \\ \vdots & \vdots \\ 1 & x_{T-1} \end{bmatrix}$$

Then

$$\begin{bmatrix} \hat{\theta} \\ \hat{\rho} \end{bmatrix} = \begin{bmatrix} \theta \\ \rho \end{bmatrix} + (X'X)^{-1} \begin{bmatrix} \sum_{t=1}^T e_t \\ \sum_{t=1}^T x_{t-1} e_t \end{bmatrix}$$

Note that given $\sigma(X) = \sigma((x_t)_{t \le T-1})$,

$$x_{t-1} \uparrow \Rightarrow e_t \downarrow$$

when $\rho > 0$. Therefore, we have

$$\mathbb{E}[x_{t-1} e_t | X] < 0$$

for $t = 1, \cdots, T-1$ and

$$\mathbb{E}\hat{\rho} < \rho$$

However, we have a consistency of $\hat{\rho}$ from

$$\begin{aligned} \lim_{T \to \infty} \begin{bmatrix} \hat{\theta} \\ \hat{\rho} \end{bmatrix} &= \begin{bmatrix} \theta \\ \rho \end{bmatrix} + (\frac{1}{T} X'X)^{-1} \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T e_t \\ \frac{1}{T} \sum_{t=1}^T x_{t-1} e_t \end{bmatrix} \\ &\to_p \begin{bmatrix} \theta \\ \rho \end{bmatrix} + (\mathbb{E}[X'X])^{-1} \begin{bmatrix} \mathbb{E} e_t \\ \mathbb{E}[x_{t-1} e_t] \end{bmatrix} \\ &= \begin{bmatrix} \theta \\ \rho \end{bmatrix} \end{aligned}$$

since

$$\begin{aligned} \mathbb{E}[X'X] &= \begin{bmatrix} 1 & \mathbb{E}x_t \\ \mathbb{E}x_t & \mathbb{E}x_t^2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \theta/(1-\rho) \\ \theta(1-\rho) & \sigma^2(1-\rho^2) + \theta^2/(1-\rho)^2 \end{bmatrix} \\ &< \infty \end{aligned}$$

and

$$\mathbb{E}[x_{t-1} e_t] = 0$$

Next, define

$$r = \begin{bmatrix} r_1 \\ \vdots \\ r_T \end{bmatrix}$$

Then

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + (X'X)^{-1} \begin{bmatrix} \sum_{t=1}^T u_t \\ \sum_{t=1}^T x_{t-1} u_t \end{bmatrix}$$

Table 8: Summary Statistics when $T = 840$

| param | n | mean | sd | skewness | kurtosis |
|---|---|---|---|---|---|
| beta | $10,000$ | 0.284 | 0.156 | 0.739 | 4.065 |
| rho | $10,000$ | 0.967 | 0.009 | -0.805 | 4.206 |

Table 9: Summary Statistics when $T = 240$

| param | n | mean | sd | skewness | kurtosis |
|---|---|---|---|---|---|
| beta | $10,000$ | 0.498 | 0.376 | 1.106 | 5.099 |
| rho | $10,000$ | 0.953 | 0.023 | -1.163 | 5.224 |

Since $\mathbb{E}[e_t u_t] < 0$, we have

$$\mathbb{E}[x_{t-1}u_t|X] > 0$$

for $t = 1, \cdots, T-1$ and

$$\mathbb{E}\hat{\beta} > \beta$$

Similarly, we have a consistency of $\hat{\beta}$ from

$$\begin{aligned}
\lim_{T\to\infty} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} &= \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} + (\frac{1}{T}X'X)^{-1} \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} u_t \\ \frac{1}{T}\sum_{t=1}^{T} x_{t-1}u_t \end{bmatrix} \\
&\to_p \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + (\mathbb{E}[X'X])^{-1} \begin{bmatrix} \mathbb{E}[u_t] \\ \mathbb{E}[x_{t-1}u_t] \end{bmatrix} \\
&= \begin{bmatrix} \alpha \\ \beta \end{bmatrix}
\end{aligned}$$

Table 8 and 9 summarise sample distributions of $\hat{\beta}$ and $\hat{\rho}$ when $T = 840$ and $T = 240$, repectively. As we expect from the finite sample properties, $\hat{\beta}$ is overestimated ($\hat{\beta} > \beta$) on average, while $\hat{\rho}$ is underestimated ($\hat{\rho} < \rho$) on average. We can also find $\hat{\beta}$ is skewed to the right (positive skewness) while $\hat{\rho}$ is skewed to the left (negative skewness). However, since both estimators are consistent, those estimation errors are quite small when sample size ($T$) is large.

6. First, note that $y_t|\varepsilon_{t-1} =_d \mathbb{N}(\mu + \theta\varepsilon_{t-1}, \sigma^2)$ for all $t$ and

$$\begin{aligned}
p_{\mu,\sigma^2,\theta}(y_t|\varepsilon_{t-1}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(y_t - \mu - \theta\varepsilon_{t-1})^2) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{\varepsilon_t^2}{2\sigma^2})
\end{aligned}$$

Second, $\sigma((y_s)_{s\leq t}, \mu, \theta) = \sigma((\varepsilon_s)_{s\leq t})$ for all $t$. Then

$$\begin{aligned}
p_{\mu,\sigma^2,\theta}(y_1, \cdots, y_t) &= p_{\mu,\sigma^2,\theta}(y_t|y_1, \cdots, y_{t-1})p_{\mu,\sigma^2,\theta}(y_1, \cdots, y_{t-1}) \\
&= p_{\mu,\sigma^2,\theta}(y_t|y_1, \cdots, y_{t-1})p_{\mu,\sigma^2,\theta}(y_{t-1}|y_1, \cdots, y_{t-2})p_{\mu,\sigma^2,\theta}(y_1, \cdots, y_{t-2}) \\
&\cdots \\
&= p_{\mu,\sigma^2,\theta}(y_t|y_1, \cdots, y_{t-1}) \times \cdots \times p_{\mu,\sigma^2,\theta}(y_2|y_1) \times p_{\mu,\sigma^2,\theta}(y_1)
\end{aligned}$$

and

$$l(y_1, \cdots, y_t, \mu, \sigma^2, \theta, ) = -\frac{T}{2}\log 2\pi - \frac{T}{2}\log\sigma^2 - \sum_{t=1}^{T}\frac{\varepsilon_t^2}{2\sigma^2}$$

where $l$ is the log likelihood function. Based on the above discussion, I conduct a Monte Carlo simulation and numerically estimate MLE. For a numerical estimation, I use a simple gradient method. Using initial parameter values $\mu = 1.9$, $\sigma^2 = 1.7$ and $\theta = 0.2$, the results from the estimation is

$$\begin{aligned}\hat{\mu}_{ML} &= 1.460,\\ \hat{\sigma}^2_{ML} &= 2.040,\\ \hat{\theta}_{ML} &= 0.0798,\end{aligned}$$

and

$$l(y, \hat{\mu}_{ML}, \hat{\sigma}^2_{ML}, \hat{\theta}_{ML}) = -1775.47$$

which is quite reasonable.

7. (a) Due to the Ito formula, we have

$$\begin{aligned}d(e^{\kappa t}(x_t - \bar{x})) &= \kappa e^{\kappa t}(x_t - \bar{x})dt + e^{\kappa t}dx_t\\ &= \kappa e^{\kappa t}(x_t - \bar{x})dt + e^{\kappa t}\kappa(\bar{x} - x_t)dt\\ &= 0\end{aligned}$$

It follows that

$$\begin{aligned}e^{\kappa t}(x_t - \bar{x}) - (x - \bar{x}) &= 0\\ x_t &= \bar{x} + e^{-\kappa t}(x - \bar{x})\end{aligned}$$

Note that

$$\begin{aligned}\lim_{t\to\infty} x_t &= \bar{x} + (x - \bar{x})\lim_{t\to\infty} e^{-\kappa t}\\ &= \bar{x}\end{aligned}$$

(b) See Figure 2.

(c) Due to the Ito formula, we have

$$\begin{aligned}d(e^{\kappa t}(x_t - \bar{x})) &= \kappa e^{\kappa t}(x_t - \bar{x})dt + e^{\kappa t}dx_t\\ &= \kappa e^{\kappa t}(x_t - \bar{x})dt + e^{\kappa t}\kappa(\bar{x} - x_t)dt + e^{\kappa t}\sigma dW_t\\ &= e^{\kappa t}\sigma dW_t\end{aligned}$$

It follows that

$$\begin{aligned}e^{\kappa t}(x_t - \bar{x}) - (x - \bar{x}) &= \sigma\int_0^t e^{\kappa s}dW_s\\ \Rightarrow x_t &= \bar{x} + e^{-\kappa t}(x - \bar{x}) + \sigma\int_0^t e^{\kappa(s-t)}dW_s.\end{aligned}$$

Then

$$\begin{aligned}\mathbb{E}_t(x_s) &= \bar{x} + e^{-\kappa s}(x - \bar{x}) + \mathbb{E}_t(\sigma\int_0^s e^{\kappa(u-s)}dW_u)\\ &= \bar{x} + e^{-\kappa t}(x - \bar{x}) + \sigma\int_0^t e^{\kappa(u-t)}dW_u + e^{-\kappa s}(x - \bar{x}) - e^{-\kappa t}(x - \bar{x})\\ &= x_t + (e^{-\kappa s} - e^{-\kappa t})(x - \bar{x}).\end{aligned}$$

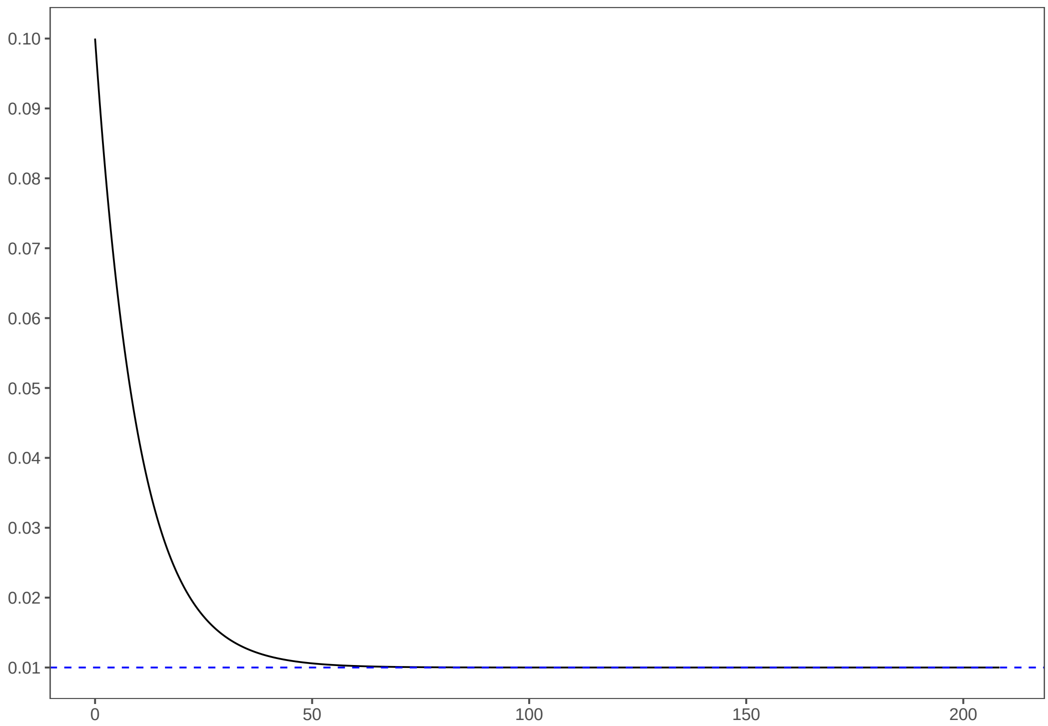


Figure 2: A discretized version of the ODE $\Delta x_t = \kappa(\bar{x} - x_t)\Delta t$. $\Delta t = 1/12$, $\kappa = 0.1$, $\bar{x} = 0.01$, $x_0 = 0.1$ and $t = 250$.

and

$$\begin{aligned}
\mathrm{var}_t(x_s) &= \mathbb{E}_t(\sigma \int_t^s e^{\kappa(u-s)} dW_u)^2 \\
&= \sigma^2 \int_t^s e^{2\kappa(u-s)} du \\
&= \sigma^2 [\frac{1}{2\kappa} e^{2\kappa(u-s)}]_t^s \\
&= \frac{\sigma^2(1-e^{2\kappa(t-s)})}{2\kappa}.
\end{aligned}$$

Note that for any given $t$,

$$\lim_{s\to\infty} \mathbb{E}_t(x_s) = x_t + (1-e^{-\kappa t})(x-\bar{x})$$

$$\begin{aligned}
\lim_{s\to\infty} \mathrm{var}_t(x_s) &= \frac{\sigma^2}{2\kappa} - \frac{1}{2\kappa} \lim_{s\to\infty} e^{2\kappa(t-s)} \\
&= \frac{\sigma^2}{2\kappa}
\end{aligned}$$

(d) See Figure 3 and 4.

# Appendix

For R codes used to produce above results, see `https://github.com/ysugk/Course-FIN625/tree/master/HW/code`.

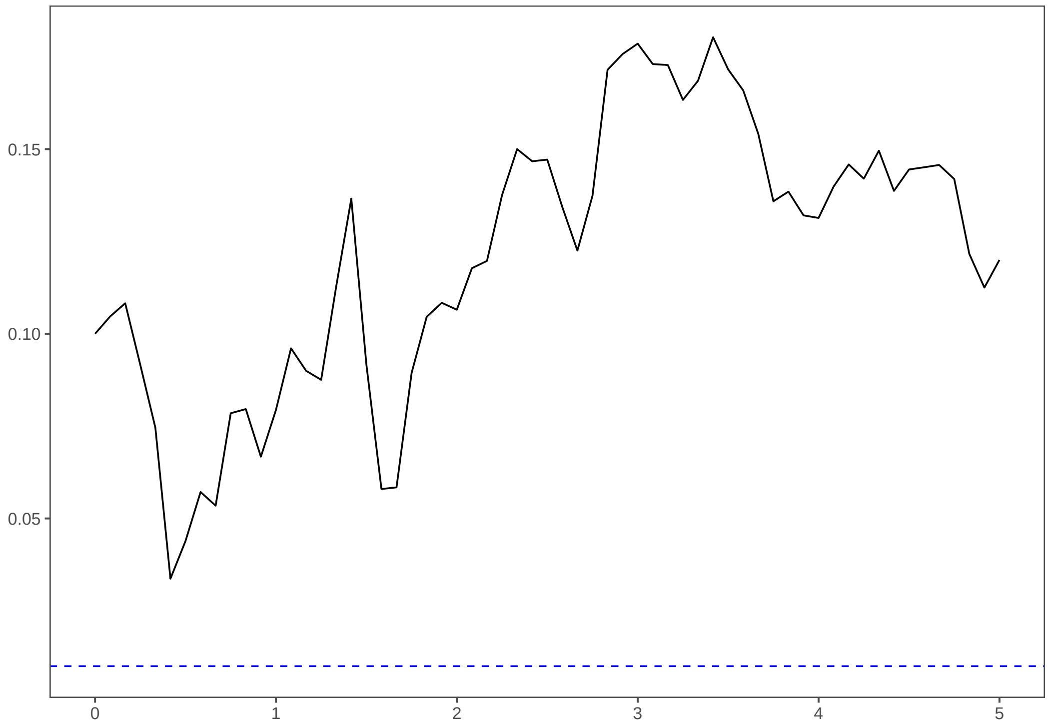


Figure 3: A discretized version of the SDE $\Delta x_t = \kappa(\bar{x} - x_t)\Delta t + \sigma(\sqrt{\Delta t})\varepsilon_t$. $\Delta t = 1/12$, $\kappa = 0.1$, $\bar{x} = 0.01$, $x_0 = 0.1$, $\sigma = 0.05$, $\varepsilon_t =_{i.i.d} \mathbb{N}(0, 1)$ and $t = 5$.


Figure 4: A discretized version of the SDE $\Delta x_t = \kappa(\bar{x} - x_t)\Delta t + \sigma(\sqrt{\Delta t})\varepsilon_t$. $\Delta t = 1/12$, $\kappa = 0.1$, $\bar{x} = 0.01$, $x_0 = 0.1$, $\sigma = 0.05$, $\varepsilon_t =_{i.i.d} \mathbb{N}(0, 1)$ and $t = 100$.