2020 秋季本科时间序列

第 6 次作业答案

12 月 10 日

1. (a) 由模型 $Y = X\beta + e$，其拟合值为 $\hat{Y} = X + \hat{\beta}$，在 OLS 估计下，$\hat{\beta} = (X^{\intercal}X)^{-1}X^{\intercal}Y$，可得

$$\hat{e} = Y - \hat{Y} = [I - X(X^{\intercal}X)^{-1}X^{\intercal}]Y = [I - X(X^{\intercal}X)^{-1}X^{\intercal}][X\beta + e] = [I - X(X^{\intercal}X)^{-1}X^{\intercal}]e$$

(b) 设 $A = [a_{ij}]_{n*k}$，$B = [b_{ij}]_{k*n}$，同时令 $C = AB = [c_{ij}]_{n*m}$，$D = BA = [d_{ij}]_{k*k}$。对于矩阵 C 上的第 $m(1 \le m \le n)$ 个对角元素 C_{mm}，易得

$$C_{mm} = \sum_{j=1}^{k} a_{mj}b_{jm}$$

进而 $\text{tr}(C) = \text{tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{k} a_{ij}b_{ji}$。同理，对于矩阵 D，可得 $\text{tr}(D) = \text{tr}(BA) = \sum_{i=1}^{k}\sum_{j=1}^{n} b_{ij}a_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{k} a_{ji}b_{ij}$，显然有 $\sum_{i=1}^{n}\sum_{j=1}^{k} b_{ij}a_{ji} = \sum_{j=1}^{n}\sum_{k=1}^{k} a_{ji}b_{ij}$（交换积分次序）。

故 $\text{tr}(C) = \text{tr}(AB) = \text{tr}(BA) = \text{tr}(D)$。若 A, B 均为 $n*n$ 的矩阵，则

$$\text{tr}(A - B) = \sum_{i=1}^{n}(a_{ii} - b_{ii}) = \sum_{i=1}^{n} a_{ii} - \sum_{i=1}^{n} b_{ii} = \text{tr}(A) - \text{tr}(B)$$

(c) 由 $\hat{e} = [\hat{e}_1, \hat{e}_2, \cdots, \hat{e}_{\intercal}]^{\intercal}$，故 $\sum_t {\hat{e}_t}^2 = (\hat{e}^{\intercal}\hat{e})_{1*1} = tr(\hat{e}^{\intercal}\hat{e})$。又由 (a) 有 $\hat{e} = [I - X(X^{\intercal}X)^{-1}X^{\intercal}]e$，故可得

$$\begin{aligned}
\hat{e}^{\intercal}\hat{e} &= [[I - X(X^{\intercal}X)^{-1}X^{\intercal}]e]^{\intercal}[[I - X(X^{\intercal}X)^{-1}X^{\intercal}]e] \\
&= e^{\intercal}[I - X(X^{\intercal}X)^{-1}X^{\intercal}]^{\intercal}[I - X(X^{\intercal}X)^{-1}X^{\intercal}]e \\
&= e^{\intercal}[I - 2X(X^{\intercal}X)^{-1}X^{\intercal} + X(X^{\intercal}X)^{-1}X^{\intercal}]e \\
&= e^{\intercal}[I - X(X^{\intercal}X)^{-1}X^{\intercal}]e
\end{aligned}$$

故 $\text{tr}(\hat{e}^{\intercal}\hat{e}) = \text{tr}(e^{\intercal}[I - X(X^{\intercal}X)^{-1}X^{\intercal}]e])$，综上得证 $\sum_t {\hat{e}_t}^2 = \text{tr}(\hat{e}^{\intercal}\hat{e}) = \text{tr}(e^{\intercal}[I - X(X^{\intercal}X)^{-1}X^{\intercal}]e])$。

(d) 由 (a),(b),(c) 问结论可知

$$\begin{aligned}\mathbb{E}[\sum_t \hat{e}_t{}^2|X] =&\mathbb{E}[\mathrm{tr}(\hat{e}^\intercal\hat{e})|X]\\ &= \mathbb{E}[\mathrm{tr}(e^\intercal[I - X(X^\intercal X)^{-1}X^\intercal]e)|X]\\ &= \mathbb{E}(\mathrm{tr}([I - X(X^\intercal X)^{-1}X^\intercal]ee^\intercal)|X)\\ &= \mathbb{E}[\mathrm{tr}(ee^\intercal)|X] - \mathbb{E}[\mathrm{tr}(X(X^\intercal X)^{-1}X^\intercal ee^\intercal)|X]\end{aligned}$$

又因为 $\mathbb{E}[\mathrm{tr}(A)] = \mathbb{E}[a_{11} + a_{22} + \cdots + a_{nn}] = \mathbb{E}(a_{11}) + \mathbb{E}(a_{22}) + \cdots + \mathbb{E}(a_{nn}) = \mathrm{tr}(\mathbb{E}[A])$。所以上式进一步简化，可得

$$\begin{aligned}&\mathbb{E}[\mathrm{tr}(ee^\intercal)|X] - \mathbb{E}[\mathrm{tr}(X(X^\intercal X)^{-1}X^\intercal ee^\intercal)|X]\\ =& \mathrm{tr}(\mathbb{E}[\mathrm{tr}(ee^\intercal)|X]) - \mathrm{tr}(\mathbb{E}[\mathrm{tr}(X(X^\intercal X)^{-1}X^\intercal ee^\intercal)|X])\\ =& \mathrm{tr}(\mathbb{E}[\mathrm{tr}(ee^\intercal)|X]) - \mathrm{tr}(\mathbb{E}[\mathrm{tr}(X(X^\intercal X)^{-1}X^\intercal \mathbb{E}(ee^T|X))\\ =& \mathrm{tr}(\sigma_e^2 I_{\intercal*\intercal}) - \mathrm{tr}((X^\intercal X)^{-1}X^\intercal\sigma_e^2)\\ =& T\sigma_e^2 - \mathrm{tr}(I_{k*k}\sigma_e^2)\\ =& T\sigma_e^2 - K\sigma_e^2 = (T-K)\sigma_e^2\end{aligned}$$

由于

$$\mathbb{E}(\frac{1}{T-K}\sum_t \hat{e}_t{}^2) = \frac{1}{T-K}\mathbb{E}[\mathbb{E}(\sum_t \hat{e}_t{}^2|X)] = \frac{1}{T-K}\mathbb{E}[(T-K)\sigma_e^2] = \sigma_e^2$$

故 $\frac{1}{T-K}\sum_t \hat{e}_t{}^2$ 是 σ_e^2 的无偏估计量。

2. (a) 对于任意 $k \in N$ 且 $k < T$，可计算 $MA(2)$ 过程 X_t 序列的自协方差为

$$\sigma^2(k) = \mathrm{cov}(X_t, X_{t-k}) = \mathbb{E}X_t X_{t-k} - 0 = \begin{cases}(1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2, & k = 0\\ (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2, & k = 1\\ \theta_2\sigma_\varepsilon^2, & k = 2\\ 0, & k \geq 3\end{cases}$$

所以 $X = [X_1, \cdots, X_\intercal]^\intercal$ 的协方差矩阵为一个五对角矩阵，如下：

$$\Sigma = \begin{bmatrix}(1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 & (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & \theta_2\sigma_\varepsilon^2 & \cdots & 0\\ (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 & (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & \cdots & \cdot\\ \theta_2\sigma_\varepsilon^2 & (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 & \cdots & \cdot\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & \cdots & \cdots & (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2\end{bmatrix}_{T*T}$$

可写出对数似然函数

$$\log L(\theta_1, \theta_2, \sigma_\varepsilon^2|X) = -\frac{1}{2}\log\det\Sigma - \frac{1}{2}X^\intercal\Sigma^{-1}X$$

理论上，通过数值方法可以求解出 $\det\Sigma$ 和 Σ^{-1}。

(b) 首先

$$|\Sigma| = \begin{vmatrix} (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 & (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & \theta_2\sigma_\varepsilon^2 & \cdots & 0 \\ (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 & (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & \cdots & \cdot \\ \theta_2\sigma_\varepsilon^2 & (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 & (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 \end{vmatrix}$$

$$= \sigma_\varepsilon^{2T} \begin{vmatrix} 1+\theta_1^2+\theta_2^2 & \theta_1+\theta_1\theta_2 & \theta_2 & \cdots & 0 \\ \theta_1+\theta_1\theta_2 & 1+\theta_1^2+\theta_2^2) & \theta_1+\theta_1\theta_2 & \cdots & \cdot \\ \theta_2 & \theta_1+\theta_1\theta_2 & 1+\theta_1^2+\theta_2^2 & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1+\theta_1^2+\theta_2^2 \end{vmatrix}$$

令 $\log L$ 分别对 $\theta_1, \theta_2, \sigma_\varepsilon^2$ 求偏导，得

$$\begin{cases} \frac{\partial \log L}{\partial \theta_1} = -\frac{1}{2\det\Sigma} * \frac{\partial \det\Sigma}{\partial \theta_1} - \frac{1}{2} * \frac{\partial X^\intercal \Sigma^{-1} X}{\partial \theta_1} = 0 \\ \frac{\partial \log L}{\partial \theta_2} = -\frac{1}{2\det\Sigma} * \frac{\partial \det\Sigma}{\partial \theta_2} - \frac{1}{2} * \frac{\partial X^\intercal \Sigma^{-1} X}{\partial \theta_2} = 0 \\ \frac{\partial \log L}{\partial \sigma_\varepsilon^2} = -\frac{T}{2} * \frac{1}{\sigma_\varepsilon^2} - \frac{1}{2} * \frac{\partial X^\intercal \Sigma^{-1} X}{\partial \sigma_\varepsilon^2} = 0 \end{cases}$$

利用数值解法可解出 $\theta_1, \theta_2, \sigma_\varepsilon^2$。

当 $\theta_2 = 0$ 时，

$$\Sigma = \sigma_\varepsilon^2 \begin{bmatrix} (1+\theta_1^2) & \theta_1 & 0 & \cdots & 0 \\ \theta_1 & (1+\theta_1^2) & \theta_1 & \cdots & \cdot \\ 0 & \theta_1 & (1+\theta_1^2) & \cdots & \cdot \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & (1+\theta_1^2) \end{bmatrix}$$

利用三对角矩阵行列式公式可得 $\det\Sigma = \frac{1-\theta_1^{2(T+1)}}{1-\theta_1^2}\sigma_\varepsilon^2$。同理，此处 Σ^{-1} 也可由数值方法求解。进而解法与上文同理。

(c) 由题意及 (a) 问可得方程组

$$\sigma^2(0) = (1+\theta_1^2+\theta_2^2)\sigma_\varepsilon^2 = a_0 \tag{1}$$

$$\sigma^2(1) = (\theta_1+\theta_1\theta_2)\sigma_\varepsilon^2 = a_1 \tag{2}$$

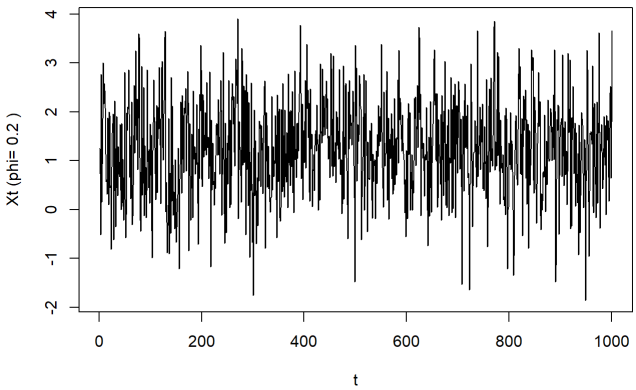
$$\sigma^2(2) = \theta_2\sigma_\varepsilon^2 = a_2 \tag{3}$$

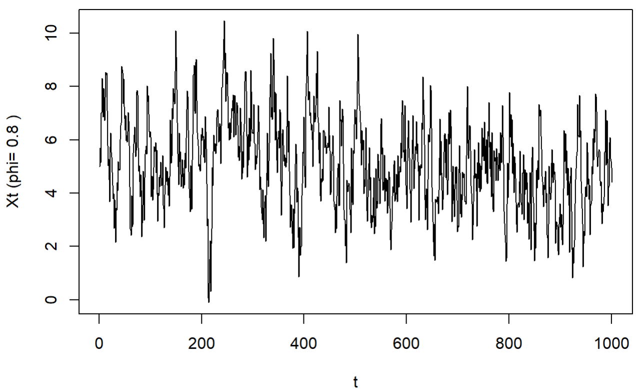
由 (3) 式可得 $\theta_2 = \frac{a_2}{\sigma_\varepsilon^2}$，代入 (2) 式，得 $\theta_1(1+\frac{a_2}{\sigma_\varepsilon^2})\frac{a_2}{\sigma_\varepsilon^2} = a$，进而 $\theta_1 = \frac{a_1}{\sigma_\varepsilon^2+a_1}$。

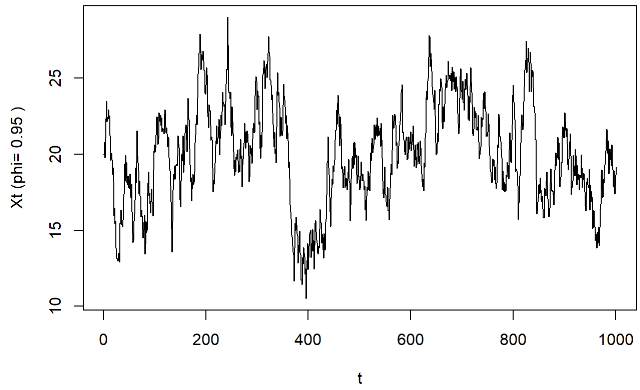
将 θ_1 和 θ_2 中关于 σ_ε^2 的表达式代入 (1) 式，得 $(1+\frac{a_1^2}{(\sigma_\varepsilon^2+a_1)^2}+\frac{a_2^2}{(\sigma_\varepsilon^2)^2})\sigma_\varepsilon^2 = a_0$，理论上可解出 σ_ε^2（四次方程）。进而代回 θ_1 和 θ_2 关于 σ_ε^2 的式子，可解出 θ_1 和 θ_2。

3. (a) 代码如下：

```r
library(tidyverse)
mu <- 1
phi <- c(0.2, 0.8, 0.95)
X <- tibble()
for(i in 1:3){
 X[1, i] = mu/(1 - phi[i])
 for(j in 1:1000){
  X[j+1, i] = mu + phi[i]*X[j, i] + rnorm(1,0,1)
 }
 plot(X[[i]], type = "l", xlab = "t", ylab = paste('Xt (phi
    =', phi[i],')'))
}
```







随着 ϕ 的增大，序列的波动性变大，但都没有呈现出明显的趋势及季节性，都是平稳时间序列。

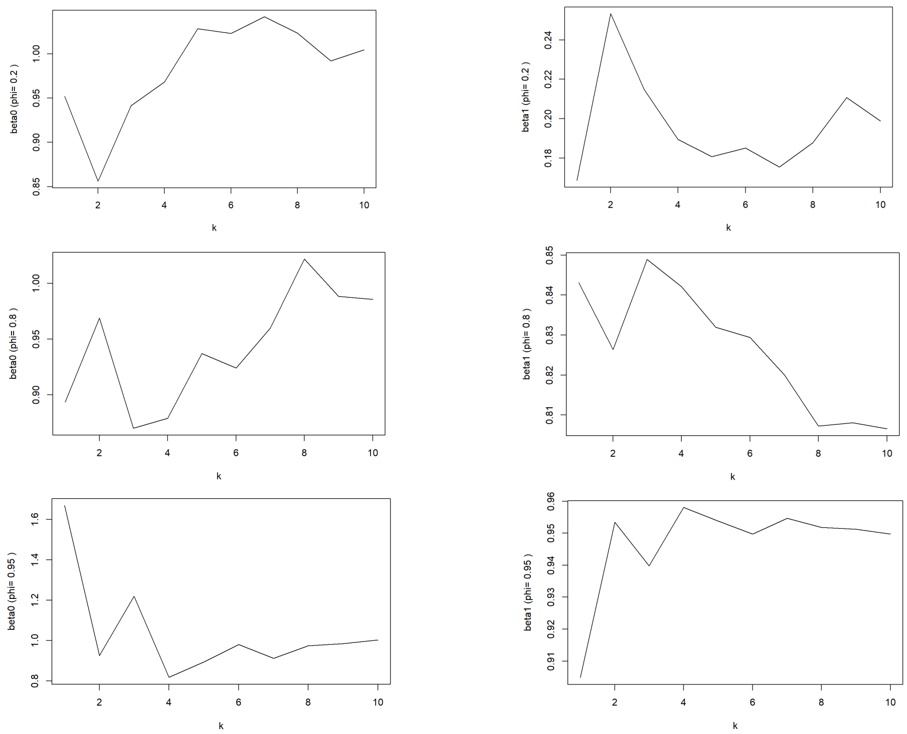
(b) 代码如下：

```
beta0 <- tibble()
beta1 <- tibble()
for(i in 1:3){
 for(k in 1:10){
  y <- unlist(X[2:(100*k+1), i])
  x <- unlist(X[1:(100*k), i])
  ols <- coefficients(lm(y ~ x))
  beta0[k, i] = ols[1]
  beta1[k, i] = ols[2]
 }
 plot(beta0[[i]], type = "l", xlab = "k", ylab = paste('
    beta0 (phi=', phi[i],')'))
 plot(beta1[[i]], type = "l", xlab = "k", ylab = paste('
    beta1 (phi=', phi[i],')'))
}
```

随着 ϕ 的增大，OLS 估计随样本量在 100k 增加时的收敛性逐渐增强。

(c) 代码如下：

```
# phi = 0.2
X1 <- tibble()
for(i in 1:500){
 X1[1, i] = mu/(1 - phi[1])
 for(j in 1:100){
  X1[j+1, i] = mu + phi[1]*X1[j, i] + rnorm(1,0,1)
```

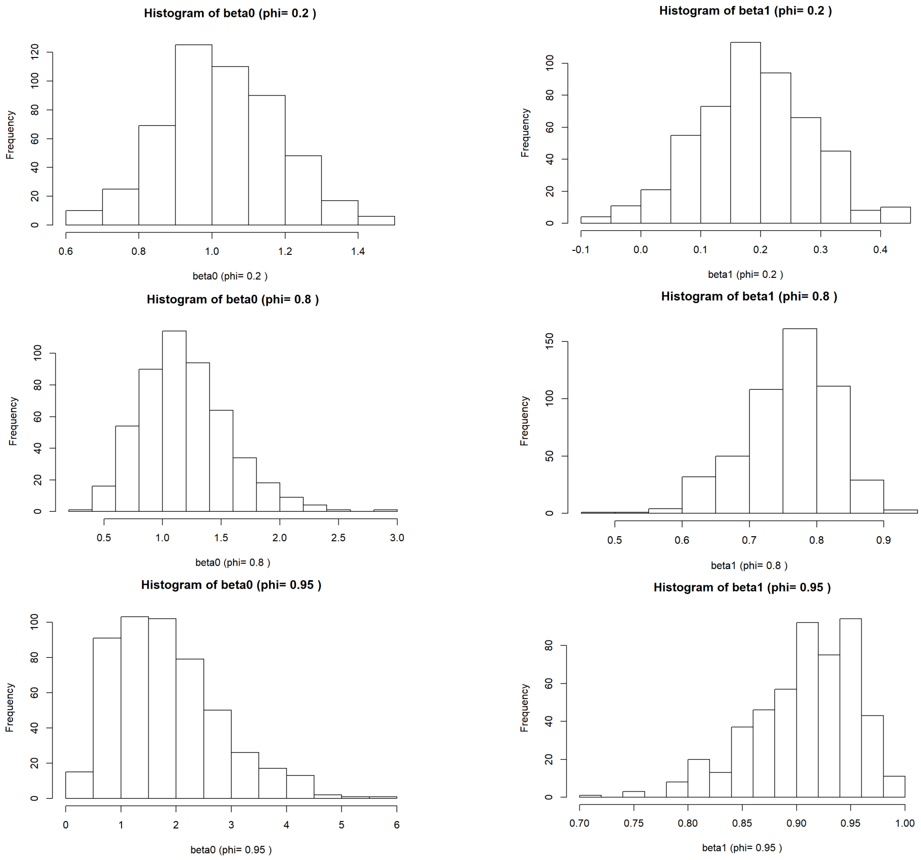



```
  }
}
# phi = 0.8
X2 <- tibble()
for(i in 1:500){
  X2[1, i] = mu/(1 - phi[2])
  for(j in 1:100){
    X2[j+1, i] = mu + phi[2]*X2[j, i] + rnorm(1,0,1)
  }
}
# phi = 0.95
X3 <- tibble()
for(i in 1:500){
  X3[1, i] = mu/(1 - phi[3])
  for(j in 1:100){
    X3[j+1, i] = mu + phi[3]*X3[j, i] + rnorm(1,0,1)
  }
}
X <- list(X1, X2, X3)
```

```
beta0 <- tibble()
beta1 <- tibble()
sd_beta0 <- vector("double",3)
sd_beta1 <- vector("double",3)
for(i in 1:3){
 for(j in 1:500){
  y <- unlist(X[[i]][2:101, j])
  x <- unlist(X[[i]][1:100, j])
  ols <- coefficients(lm(y ~ x))
  beta0[j, i] <- ols[1]
  beta1[j, i] <- ols[2]
 }
 hist(beta0[[i]], xlab = paste('beta0 (phi=', phi[i],')'),
    main = paste('Histogram of beta0 (phi=', phi[i],')'))
 hist(beta1[[i]], xlab = paste('beta1 (phi=', phi[i],')'),
    main = paste('Histogram of beta1 (phi=', phi[i],')'))
 sd_beta0[i] <- sd(beta0[[i]])
 sd_beta1[i] <- sd(beta1[[i]])
}
sd_beta0
## [1] 0.1581694 0.3707317 0.9623762
sd_beta1
## [1] 0.09468137 0.06826351 0.04768194
```

(d) 代码如下：

```
 x <- vector("double",101)
 sd_beta0 <- vector("double",3)
 sd_beta1 <- vector("double",3)
 for (i in 1:3) {
  x[1] <- mu/(1 - phi[i])
  for (j in 1:100) {
   x[j+1] <- mu + phi[i]*x[j] + rnorm(1,0,1)
  }
  X <- cbind(rep(1,100), x[2:101])
  M <- t(X) %*% X / 100   # 求M矩阵
```




```
  sd_beta0[i] <- sqrt(1/100*1*solve(M)[1,1])
  sd_beta1[i] <- sqrt(1/100*1*solve(M)[2,2])
 }
 sd_beta0
 ## [1] 0.1515409 0.3606424 1.0093074
 sd_beta1
 ## [1] 0.09815747 0.07496892 0.04951775
```

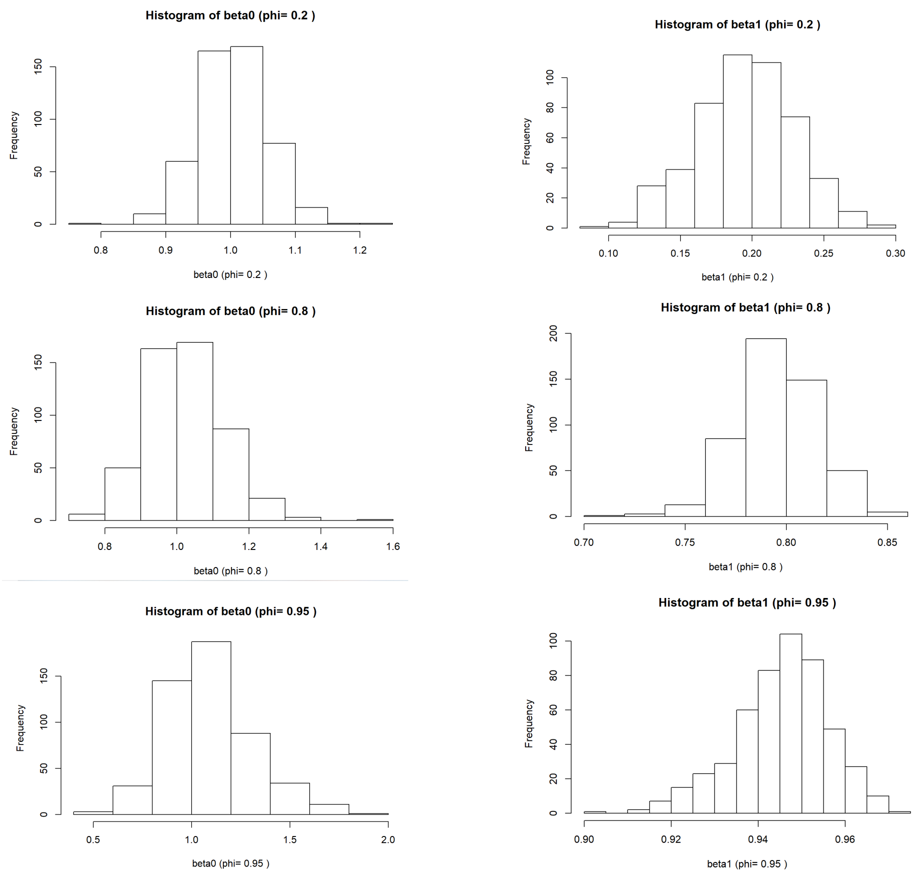
渐进标准误接近样本标准差的值。

(e) 代码如下:

```
# phi = 0.2
X1 <- tibble()
for(i in 1:500){
 X1[1, i] = mu/(1 - phi[1])
 for(j in 1:900){
  X1[j+1, i] = mu + phi[1]*X1[j, i] + rnorm(1,0,1)
 }
}
```

```
# phi = 0.8
X2 <- tibble()
for(i in 1:500){
 X2[1, i] = mu/(1 - phi[2])
 for(j in 1:900){
  X2[j+1, i] = mu + phi[2]*X2[j, i] + rnorm(1,0,1)
 }
}
# phi = 0.95
X3 <- tibble()
for(i in 1:500){
 X3[1, i] = mu/(1 - phi[3])
 for(j in 1:900){
  X3[j+1, i] = mu + phi[3]*X3[j, i] + rnorm(1,0,1)
 }
}
X <- list(X1, X2, X3)

beta0 <- tibble()
beta1 <- tibble()
sd_beta0 <- vector("double",3)
sd_beta1 <- vector("double",3)
for(i in 1:3){
 for(j in 1:500){
  y <- unlist(X[[i]][2:901, j])
  x <- unlist(X[[i]][1:900, j])
  ols <- coefficients(lm(y ~ x))
  beta0[j, i] <- ols[1]
  beta1[j, i] <- ols[2]
 }
 hist(beta0[[i]], xlab = paste('beta0 (phi=', phi[i],')'),
    main = paste('Histogram of beta0 (phi=', phi[i],')'))
 hist(beta1[[i]], xlab = paste('beta1 (phi=', phi[i],')'),
    main = paste('Histogram of beta1 (phi=', phi[i],')'))
 sd_beta0[i] <- sd(beta0[[i]])
```




```
  sd_beta1[i] <- sd(beta1[[i]])
}
```

此时估计系数的样本标准差近似为 (c) 中的 1/3。