STAT5044_HW6

zhengzhi lin

2019.11.8

```r
t <- as.data.frame(matrix(0,nrow = 12,ncol = 8))
colnames(t) <- c("case","method","bias_beta1",
                 "var_beta1","mse_beta1","bias_beta2","var_beta2","mse_beta2")

set.seed(500)
#case 1
library(L1pack)
library(MASS)
library(knitr)

lse_es <- matrix(0,nrow = 100,ncol = 2)
lad_es <- matrix(0,nrow = 100,ncol = 2)
huber_es <- matrix(0,nrow = 100,ncol = 2)
lts_es <- matrix(0,nrow = 100,ncol = 2)

beta1 <- 1
beta2 <- 2
beta <- c(1,2)
for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  lse <- lm(Y~X)
  lse_es[i,] <- lse$coefficients
}
lse_mse <- diag(var(lse_es)) + (beta-mean(lse_es))^2
t[1,] <- c(1,"lse",(beta-mean(lse_es))[1],diag(var(lse_es))[1],
           lse_mse[1],(beta-mean(lse_es))[2],diag(var(lse_es))[2],lse_mse[2])

for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  Lad <- lad(Y~X)
  lad_es[i,] <- Lad$coefficients
}
lad_mse <- diag(var(lad_es)) + (beta-mean(lad_es))^2
t[2,] <- c(1,"lad",(beta-mean(lad_es))[1],diag(var(lad_es))[1],
           lad_mse[1],(beta-mean(lad_es))[2],diag(var(lad_es))[2],lad_mse[2])


for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  huber<- rlm(Y~X)
  huber_es[i,] <- huber$coefficients
```

```
}
huber_mse <- diag(var(huber_es)) + (beta-mean(huber_es))^2
t[3,] <- c(1,"huber",(beta-mean(huber_es))[1],diag(var(huber_es))[1],
           huber_mse[1],(beta-mean(huber_es))[2],diag(var(huber_es))[2],huber_mse[2])


for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  lts<- ltsreg(Y~X)
  lts_es[i,] <- lts$coefficients
}

lts_mse <- diag(var(lts_es)) + (beta-mean(lts_es))^2
t[4,] <- c(1,"lts",(beta-mean(lts_es))[1],diag(var(lts_es))[1],
           lts_mse[1],(beta-mean(lts_es))[2],diag(var(lts_es))[2],lts_mse[2])

#case2
lse_es <- matrix(0,nrow = 100,ncol = 2)
lad_es <- matrix(0,nrow = 100,ncol = 2)
huber_es <- matrix(0,nrow = 100,ncol = 2)
lts_es <- matrix(0,nrow = 100,ncol = 2)

for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rlnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  lse <- lm(Y~X)
  lse_es[i,] <- lse$coefficients
}
lse_mse <- diag(var(lse_es)) + (beta-mean(lse_es))^2
t[5,] <- c(2,"lse",(beta-mean(lse_es))[1],diag(var(lse_es))[1],
           lse_mse[1],(beta-mean(lse_es))[2],diag(var(lse_es))[2],lse_mse[2])

for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rlnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  Lad <- lad(Y~X)
  lad_es[i,] <- Lad$coefficients
}
lad_mse <- diag(var(lad_es)) + (beta-mean(lad_es))^2
t[6,] <- c(2,"lad",(beta-mean(lad_es))[1],diag(var(lad_es))[1],
           lad_mse[1],(beta-mean(lad_es))[2],diag(var(lad_es))[2],lad_mse[2])


for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rlnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  huber<- rlm(Y~X)
  huber_es[i,] <- huber$coefficients
```

```r
}
huber_mse <- diag(var(huber_es)) + (beta-mean(huber_es))^2
t[7,] <- c(2,"huber",(beta-mean(huber_es))[1],diag(var(huber_es))[1],
           huber_mse[1],(beta-mean(huber_es))[2],diag(var(huber_es))[2],huber_mse[2])


for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rlnorm(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  lts<- ltsreg(Y~X)
  lts_es[i,] <- lts$coefficients
}

lts_mse <- diag(var(lts_es)) + (beta-mean(lts_es))^2
t[8,] <- c(2,"lts",(beta-mean(lts_es))[1],diag(var(lts_es))[1],
           lts_mse[1],(beta-mean(lts_es))[2],diag(var(lts_es))[2],lts_mse[2])

#case3
lse_es <- matrix(0,nrow = 100,ncol = 2)
lad_es <- matrix(0,nrow = 100,ncol = 2)
huber_es <- matrix(0,nrow = 100,ncol = 2)
lts_es <- matrix(0,nrow = 100,ncol = 2)

for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rcauchy(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  lse <- lm(Y~X)
  lse_es[i,] <- lse$coefficients
}
lse_mse <- diag(var(lse_es)) + (beta-mean(lse_es))^2
t[9,] <- c(3,"lse",(beta-mean(lse_es))[1],diag(var(lse_es))[1],
           lse_mse[1],(beta-mean(lse_es))[2],diag(var(lse_es))[2],lse_mse[2])

for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rcauchy(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  Lad <- lad(Y~X)
  lad_es[i,] <- Lad$coefficients
}
lad_mse <- diag(var(lad_es)) + (beta-mean(lad_es))^2
t[10,] <- c(3,"lad",(beta-mean(lad_es))[1],diag(var(lad_es))[1],
            lad_mse[1],(beta-mean(lad_es))[2],diag(var(lad_es))[2],lad_mse[2])


for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rcauchy(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  huber<- rlm(Y~X)
  huber_es[i,] <- huber$coefficients
```

```
}
huber_mse <- diag(var(huber_es)) + (beta-mean(huber_es))^2
t[11,] <- c(3,"huber",(beta-mean(huber_es))[1],diag(var(huber_es))[1],
            huber_mse[1],(beta-mean(huber_es))[2],diag(var(huber_es))[2],huber_mse[2])


for (i in 1:100) {
  X <- rnorm(100,0,1)
  epsilon <- rcauchy(100,0,2)
  Y <- beta1 + beta2*X + epsilon
  lts<- ltsreg(Y~X)
  lts_es[i,] <- lts$coefficients
}

lts_mse <- diag(var(lts_es)) + (beta-mean(lts_es))^2
t[12,] <- c(3,"lts",(beta-mean(lts_es))[1],diag(var(lts_es))[1],
            lts_mse[1],(beta-mean(lts_es))[2],diag(var(lts_es))[2],lts_mse[2])

for (i in 3:8) {
  t[,i] <- round(as.numeric(t[,i]),2)
}
t$mse_sum <- t[,5] + t[,8]
kable(t)
```

case	method	bias_beta1	var_beta1	mse_beta1	bias_beta2	var_beta2	mse_beta2	mse_sum
1	lse	-0.48	0.04	0.27	0.52	0.04	0.31	0.58
1	lad	-0.49	0.06	0.30	0.51	0.07	0.32	0.62
1	huber	-0.51	0.04	0.30	0.49	0.04	0.28	0.58
1	lts	-0.49	0.31	0.55	0.51	0.29	0.55	1.10
2	lse	-4.70	23.04	45.15	-3.70	48.20	61.91	107.06
2	lad	-1.07	0.09	1.23	-0.07	0.08	0.09	1.32
2	huber	-1.50	0.22	2.47	-0.50	0.09	0.35	2.82
2	lts	-0.69	0.01	0.48	0.31	0.00	0.10	0.58
3	lse	-1.08	991.52	992.70	-0.08	1977.57	1977.58	2970.28
3	lad	-0.47	0.09	0.31	0.53	0.10	0.39	0.70
3	huber	-0.47	0.20	0.42	0.53	0.19	0.47	0.89
3	lts	-0.52	0.20	0.47	0.48	0.18	0.41	0.88

As we see in the table. For case 1, the best method is LSE because it satisfies all assumptions of LSE. For case 2 and 3 error is not normal thus LSE performs bad. In case 2, the best method is LTS, because the lognormal error has a lot extreme value, LTS is less sensitive to outliers than other methods. In case 3, the best method is LAD,