ARE212: Section 04

Dan Hammer

August 13, 2012

This is an introducton to basic hypothesis testing in R. We have shown that, with a certain set of assumptions, the difference between the OLS estimator and the true parameter vector is distributed normally as shown in expression (2.63):

$$(\mathbf{b} - \beta)|\mathbf{X} \sim N(\mathbf{0},\ \sigma^2 \cdot (\mathbf{X}'\mathbf{X})^{-1})$$

We have also shown that $s^2 = \mathbf{e}'\mathbf{e}/(n-k)$ is an unbiased estimator of σ^2 in Section 2.3.4 of the lecture notes. The purpose of the section is not to rehash the lectures, but instead to use the results to practice indexing in R.

```
data <- read.csv("../data/auto.csv", header=TRUE)
names(data) <- c("price", "mpg", "weight")
y <- matrix(data$price)
X <- cbind(1, data$mpg, data$weight)

n <- nrow(X); k <- ncol(X)
P <- X %*% solve(t(X) %*% X) %*% t(X)
e <- (diag(n) - P) %*% y
s2 <- t(e) %*% e / (n - k)
print(s2)

         [,1]
 [1,] 6320340

vcov.mat <- as.numeric(s2) * solve(t(X) %*% X)
se.vec <- diag(vcov.mat)
se.vec

 [1] 1.293877e+07 7.422863e+03 4.113347e-01
```