Aula 10: Testes de Robustez

**Lista de Exercícios**

07 & 09/11/2022

1. Discuta os conceitos de validade interna e externa.
2. Quais são as ameaças à validade interna na regressão múltipla?
3. As simulações a seguir foram feitas com base em “Introduction to Econometrics with R” (HANCK et al, 2020). Estimaremos dois modelos de regressão 1000 vezes cada um e depois obteremos a média dos coeficientes. Compararemos esses valores com os parâmetros da população. Em cada caso, identifique qual é a ameaça à validade interna nos dois modelos.

**Modelo 1**

Nesta simulação a função de regressão real da população é:

Se estimarmos a regressão abaixo encontraremos o da população? Compare o com o da população, que é igual a 0.

Podemos encontrar o mesmo beta\_0 por coincidência, então, estimaremos a regressão 1000 vezes e depois comparemos a médias dos estimados com o da população. Por que estamos estimando a regressão 1000 vezes ao invés de apenas uma vez?

Características dos dados simulados:

*n = 100*X ∼ U[-5,5]: distribuição uniforme que varia entre -5 e 5 \* u ~ N(0, 1).

*#definindo o seed*

set.seed(42)

*#defina o numero de iteracoes*

N <- 1000

*#crie um vetor dos beta\_hats*

beta\_0 <- c()

*#estimando os 1000 beta\_hats*

*##loop: pedindo para rodar o modelo N(1k) vezes*

**for** (i **in** 1:N) {

*#simulando os dados*

X <- runif(100, -5, 5)

Y <- X^2 + rnorm(100)

*#estimando a regressao linear*

ms\_mod <- lm(Y ~X)

*#salvando as estimativas*

beta\_0[i] <- ms\_mod$coefficients[1]

}

*#obtendo a media dos beta\_0 estimados*

mean(beta\_0)

## [1] 8.30133

*# esse valor eh igual a o?*

mean(beta\_0) == 0

## [1] FALSE

**Modelo 2**

A função de regressão real da população é:

Os parâmetros da regressão são:

Quais são os valores do e da população?

Estimaremos a função de regressão abaixo, compare com a da população e discuta o que estamos mensurando errado. O será igual ao da população?

Instruções:

Gere 1000 estimativas de através do modelo abaixo:

Salve os coeficientes como e ;

Obtenha a média dos coeficientes e compare com o da população.

*# instalando o pacote*

install.packages("mvtnorm")

*# definindo o seed*

set.seed(42)

*# carregando o pacote mvtnorm*

**library**(mvtnorm)

*# definindo o numero de iteracoes*

N <- 1000

*# vetor de beta\_1*

beta\_1 <- c()

*# definindo o loop*

**for** (i **in** 1:N) {

*# simulando o dataset e os nomes das colunas*

d <- data.frame(rmvnorm(1000, c(50, 100), sigma = cbind(c(10, 5), c(5, 10))))

colnames(d) <- c("X", "Y")

*# adicionando o erro a mensuracao de x*

d[,1] <- d[,1] + rnorm(100,0,sqrt(10))

*# estimando a regressao linear simples*

ms\_mod <- lm(Y ~ X, data = d)

*# salvando o beta\_1*

beta\_1[i] <- ms\_mod$coefficients[2]

}

*# obtendo a media dos beta\_1 estimados*

mean(beta\_1)

## [1] 0.2526042

*# a media dos betas\_1 eh igual a 0.5*

mean(beta\_1) == 0.5

## [1] FALSE

Referências

Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2019). Introduction to Econometrics with R. Obtenido de https://www. econometrics-with-r. org/ITER. pdf.