Exercício semana 8

Vinícius Araújo Kluska - GRR20183601

26/06/2021

# 1) Regressão Múltipla com Mínimos Quadrados Ordinários (MQO)

Utilizando os dados da tabela “crime1.txt” extraída do Wooldridge e disponível no Moodle, estime a regressão de narr86 como variável dependente em função de pcnv, avgsen, ptime86, qemp86. A descrição das variáveis se encontra no arquivo texto “crime1.DES”, também disponível no Moodle.

* 1. Apresente os resultados obtidos com o R incluindo comandos, estimativas, teste t, teste F e R quadrado e analise-os.
  2. Analise o teste t. Apresente a H0 e responda à questão: os coeficientes da regressão são significativos a 5%?
  3. Analise o teste F. Apresente a H0 e responda à questão: os coeficientes da regressão são significativos conjuntamente?
  4. Analise o R quadrado, respondendo à questão: quanto a regressão explica das variações da variável dependente?

## Resolução:

Open data base “Crime1”

tCrime<-read.delim2("C:/Users/klusk/Projects/R-studies/Econometrics\_S308/Crime\_1.txt")

Establish regression

regCrime=lm(narr86~pcnv+avgsen+ptime86+qemp86, data=tCrime)  
Regression<-summary(regCrime)  
print(Regression)

##   
## Call:  
## lm(formula = narr86 ~ pcnv + avgsen + ptime86 + qemp86, data = tCrime)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.9330 -0.4247 -0.2934 0.3506 11.4403   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.706756 0.033151 21.319 < 2e-16 \*\*\*  
## pcnv -0.150832 0.040858 -3.692 0.000227 \*\*\*  
## avgsen 0.007443 0.004734 1.572 0.115993   
## ptime86 -0.037391 0.008794 -4.252 2.19e-05 \*\*\*  
## qemp86 -0.103341 0.010396 -9.940 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.8414 on 2720 degrees of freedom  
## Multiple R-squared: 0.04219, Adjusted R-squared: 0.04079   
## F-statistic: 29.96 on 4 and 2720 DF, p-value: < 2.2e-16

## Análise de resultados:

* 1. A regressão é estimada da seguinte forma:
  + narr86 = 0,7067456 - (0,150832pcnv) + (0,007443avgsen) - (0,037391ptime86) -(0,103341qemp86)
  1. Com relação aos betas, somente “avgsen” não reijeita H0 a 5% de significância pois é o único com p-valor superior a 0,05.
  2. Sobre a regressão como um todo, o resultado do teste F demonstra que esta é significante a 5%.
  3. Segundo o R², a regressão estimada explica 4,219% da variação, o que é considerado um baixo poder explicativo. O R² ajustado apresenta um ajuste de 4,079%.

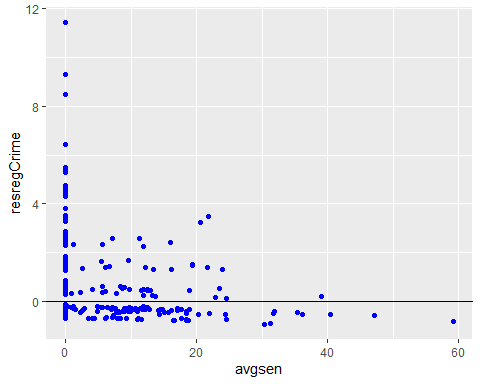
# 2) Gráfico dos resíduos da regressão estimada

Desenhe o gráfico dos resíduos da regressão para observar sua dispersão (plotar a E(u) = 0 é recomendável para facilitar a observação). Observe o gráfico e analise se: a) A variância do resíduo parece constante, ou seja, homocedástica; b) Os resíduos parecem seguir a distribuição normal (simetria, por exemplo)

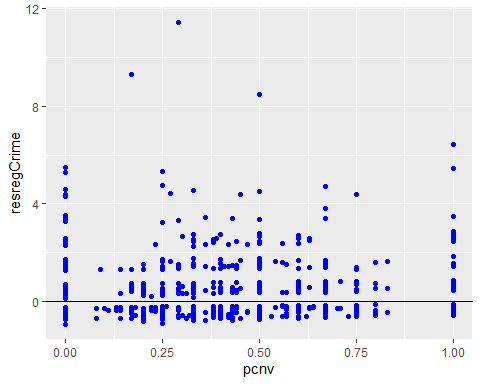
## Resolução:

Plot Graph: Residuals

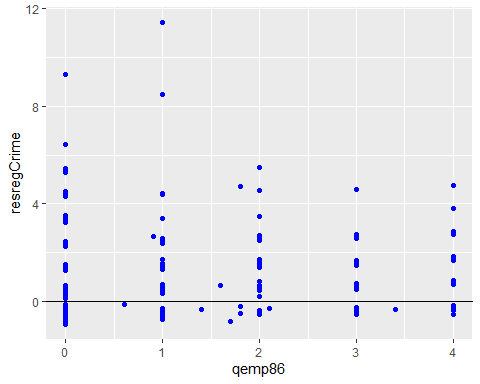
resregCrime <- regCrime$residuals  
library(ggplot2)  
Gresavgsen<-ggplot(data = tCrime, aes(y=resregCrime, x = avgsen))+geom\_point(col='Blue')+geom\_abline(slope = 0)  
Grespcnv<-ggplot(data = tCrime, aes(y=resregCrime, x = pcnv))+geom\_point(col='Blue')+geom\_abline(slope = 0)  
Gresptime86<-ggplot(data = tCrime, aes(y=resregCrime, x = ptime86))+geom\_point(col='Blue')+geom\_abline(slope = 0)  
Gresqemp86<-ggplot(data = tCrime, aes(y=resregCrime, x = qemp86))+geom\_point(col='Blue')+geom\_abline(slope = 0)  
print(Gresavgsen)



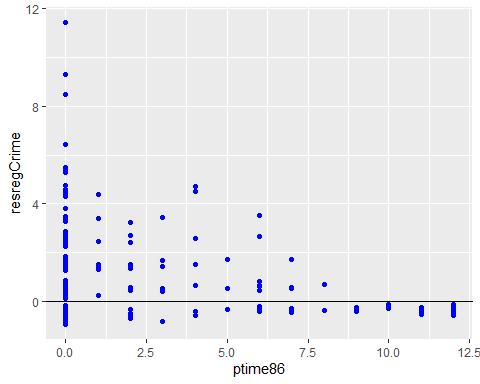
print(Grespcnv)



print(Gresqemp86)



print(Gresptime86)



## Análise de resultados:

* O gráfico dos resíduos da regressão com relação a variável “avgsen” apresenta aparente heterocedasticidade e assimetria com relação a distribuição normal.
* O gráfico dos resíduos da regressão com relação a variável “pcnv” apresenta aparente homocedasticidade e simetria para com relação a distribuição normal.
* O gráfico dos resíduos da regressão com relação a variável “qemp86” apresenta aparente homocedasticidade e simetria para com relação a distribuição normal.
* O gráfico dos resíduos da regressão com relação a variável “ptime86” apresenta aparente heterocedasticidade e assimetria com relação a distribuição normal.

Para uma análise mais embasada um teste BPG para a avaliação indivídual dos parâmetros - H0: homocedasticidade e H1: heterocedasticidade

# 3) Teste de White com e sem termos cruzados

Aplique o teste de White à regressão com e sem termos cruzados (é necessário o pacote “skedastic”). Apresente a H0 e analise se eles apontam ou não violação da hipótese de homocedasticidade.

## Resolução:

Run White\_test:

library(skedastic)  
White\_test=white\_lm(regCrime, interactions=FALSE, statonly=FALSE)  
White\_test\_crossed=white\_lm(regCrime, interactions=TRUE, statonly=FALSE)  
print(White\_test)

## # A tibble: 1 x 5  
## statistic p.value parameter method alternative  
## <dbl> <dbl> <dbl> <chr> <chr>   
## 1 58.4 9.64e-10 8 White's Test greater

print(White\_test\_crossed)

## # A tibble: 1 x 5  
## statistic p.value parameter method alternative  
## <dbl> <dbl> <dbl> <chr> <chr>   
## 1 59.1 0.000000169 14 White's Test greater

## Análise de resultados:

* 1. H0: homocedasticidade / H1: heterocedasticidade.
  2. Teste White sem interações:
  + P-value > 5%, logo rejeitar H0. Teste White com interações:
  + P-value > 5%, logo rejeitar H0. Resumo: Indício de heterocedasticidade.

# 4) Teste t robusto

Refaça os testes t utilizando a variância robusta de white (os pacotes lmtest e sandwich são necessários). Apresente a H0 e analise se o resultado do teste indica que os coeficientes são significativos a 5%.

## Resolução:

Run T\_robust test

library(lmtest)  
library(sandwich)  
T\_robust=coeftest(regCrime, vcov=vcovHC(regCrime, type = 'HC0'))  
print(T\_robust)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.7067565 0.0411344 17.1817 < 2.2e-16 \*\*\*  
## pcnv -0.1508319 0.0338607 -4.4545 8.749e-06 \*\*\*  
## avgsen 0.0074431 0.0052028 1.4306 0.1527   
## ptime86 -0.0373908 0.0061183 -6.1113 1.129e-09 \*\*\*  
## qemp86 -0.1033410 0.0115564 -8.9423 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Análise de resultados:

* H0: O estimador é significante a 5% de significância / H1: O estimador não é significante a 5% de significância.
  + A variância de White é uma forma de conferir maior peso para as informações que apresentam menor variância (informações mais próximas a média populacional).
* O teste mantem que avgsen é não significativo a 5%.