

Variables de control - Agentes libres

Antonio Huerta Montellano

15 de junio del 2022

Importando las librerías:

Veamos cuál es el directorio de trabajo

```
getwd()
```

```
## [1] "/home/usuario/Documentos/Github/Proyectos/MLB_HN/Models/Instrumental_variables/Free_agent"
```

Cambiamos el directorio de trabajo y carguemos las bases de datos para el modelo lineal en el mismo chunk:

```
setwd("~/Documentos/Github/Proyectos/MLB_HN/")
free_agents <- read.csv('Data/New_Data/Models/Article/article_iv_linear_regression_fa.csv')
```

Observemos el contenido de las bases de datos de los agentes libres:

```
head(free_agents)
```

##	Jugador	Y	X	Posicion	Equipo	Equipos_estado
## 1	A.J. Burnett	-471.0402	0.6333714	SP	PHI	2
## 2	A.J. Pierzynski	199.1753	-0.5233840	C	ATL	1
## 3	A.J. Pierzynski	1218.5682	-0.6992366	C	STL	2
## 4	A.J. Pierzynski	-2515.7319	0.7312696	C	TEX	2
## 5	Aaron Harang	509.2721	-0.6291943	SP	LAD	5
## 6	Adam LaRoche	293.4260	0.6825397	DH	WSH	1

Dummificación

Creemos las variables dummy correspondientes a las variables categóricas de la posición del jugador y al estad

```
dummy <- dummy_cols(free_agents, select_columns = c('Posicion', 'Equipo'),
                     remove_selected_columns = TRUE)
```

Notemos que ahora hay muchas columnas debido a la dummificación y la gran cantidad de categorías tanto para la posición que puede ocupar un jugador como para el equipo al que pertenecen (29):

```
head(dummy)
```

```
##          Jugador          Y          X Equipos_estado Posicion_1B Posicion_2B
## 1    A.J. Burnett -471.0402  0.6333714          2          0          0
## 2  A.J. Pierzynski  199.1753 -0.5233840          1          0          0
## 3  A.J. Pierzynski 1218.5682 -0.6992366          2          0          0
## 4  A.J. Pierzynski -2515.7319  0.7312696          2          0          0
## 5    Aaron Harang  509.2721 -0.6291943          5          0          0
## 6    Adam LaRoche  293.4260  0.6825397          1          0          0
##  Posicion_3B Posicion_C Posicion_CF Posicion_DH Posicion_LF Posicion_RF
## 1          0          0          0          0          0          0
## 2          0          1          0          0          0          0
## 3          0          1          0          0          0          0
## 4          0          1          0          0          0          0
## 5          0          0          0          0          0          0
## 6          0          0          0          1          0          0
##  Posicion_RP Posicion_RP/CL Posicion_SP Posicion_SS Equipo_ARI Equipo_ATL
## 1          0          0          1          0          0          0
## 2          0          0          0          0          0          1
## 3          0          0          0          0          0          0
## 4          0          0          0          0          0          0
## 5          0          0          1          0          0          0
## 6          0          0          0          0          0          0
##  Equipo_BAL Equipo_BOS Equipo_CHC Equipo_CHW Equipo_CIN Equipo_COL Equipo_DET
## 1          0          0          0          0          0          0          0
## 2          0          0          0          0          0          0          0
## 3          0          0          0          0          0          0          0
## 4          0          0          0          0          0          0          0
## 5          0          0          0          0          0          0          0
## 6          0          0          0          0          0          0          0
##  Equipo_HOU Equipo_KC Equipo_LAA Equipo_LAD Equipo_MIA Equipo_MIL Equipo_MIN
## 1          0          0          0          0          0          0          0
## 2          0          0          0          0          0          0          0
## 3          0          0          0          0          0          0          0
## 4          0          0          0          0          0          0          0
## 5          0          0          0          1          0          0          0
## 6          0          0          0          0          0          0          0
##  Equipo_NYM Equipo_NYY Equipo_OAK Equipo_PHI Equipo_PIT Equipo_SD Equipo_SEA
## 1          0          0          0          1          0          0          0
## 2          0          0          0          0          0          0          0
## 3          0          0          0          0          0          0          0
## 4          0          0          0          0          0          0          0
## 5          0          0          0          0          0          0          0
## 6          0          0          0          0          0          0          0
##  Equipo_SF Equipo_STL Equipo_TB Equipo_TEX Equipo_TOR Equipo_WSH
## 1          0          0          0          0          0          0
## 2          0          0          0          0          0          0
## 3          0          1          0          0          0          0
## 4          0          0          0          1          0          0
## 5          0          0          0          0          0          0
## 6          0          0          0          0          0          1
```

Creación del modelo

Para ello y evitar errores, lo haremos con los nombres explícitos, obtengamos los nombres de las columnas

```
colnames(dummy)
```

```
## [1] "Jugador"      "Y"            "X"            "Equipos_estado"
## [5] "Posicion_1B"   "Posicion_2B"   "Posicion_3B"   "Posicion_C"
## [9] "Posicion_CF"   "Posicion_DH"   "Posicion_LF"   "Posicion_RF"
## [13] "Posicion_RP"   "Posicion_RP/CL" "Posicion_SP"   "Posicion_SS"
## [17] "Equipo_ARI"    "Equipo_ATL"    "Equipo_BAL"    "Equipo_BOS"
## [21] "Equipo_CHC"    "Equipo_CHW"    "Equipo_CIN"    "Equipo_COL"
## [25] "Equipo_DET"    "Equipo_HOU"    "Equipo_KC"     "Equipo_LAA"
## [29] "Equipo_LAD"    "Equipo_MIA"    "Equipo_MIL"    "Equipo_MIN"
## [33] "Equipo_NYM"    "Equipo_NYY"    "Equipo_OAK"    "Equipo_PHI"
## [37] "Equipo_PIT"    "Equipo_SD"     "Equipo_SEA"    "Equipo_SF"
## [41] "Equipo_STL"    "Equipo_TB"     "Equipo_TEX"    "Equipo_TOR"
## [45] "Equipo_WSH"
```

Cambiamos el nombre de la columna “Posicion_RP/CL” por “Posicion_RP_CL” para evitar problemas en los algoritmos:

```
names(dummy)[names(dummy) == 'Posicion_RP/CL'] <- 'Posicion_RP_CL'
```

```
formula_iv <- Y ~ X | Equipos_estado + Posicion_1B + Posicion_2B + Posicion_3B + Posicion_C + Posicion_
iv_model <- ivreg(formula_iv, data = dummy)
summary(iv_model, diagnostics=TRUE)
```

```
##
## Call:
## ivreg(formula = formula_iv, data = dummy)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3462.88  -562.46   89.31   593.83  2908.09
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -158.47     68.33  -2.319  0.0209 *
## X              117.91    168.14   0.701  0.4836
##
## Diagnostic tests:
##              df1 df2 statistic    p-value
## Weak instruments  39 323     2.724 0.000000773 ***
## Wu-Hausman        1 360     0.433     0.511
## Sargan            41  NA    44.823     0.315
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1015 on 361 degrees of freedom
## Multiple R-Squared: -0.003463, Adjusted R-squared: -0.006243
## Wald test: 0.4917 on 1 and 361 DF, p-value: 0.4836
```

En comparación con el modelo sin instrumentos -donde se obtuvo un $p - value = 0.738$ -, aquí mejoró dicho valor. Sin embargo, no es suficiente para que la variable X sea estadísticamente significativa para el modelo. También dio positivo para la prueba de instrumentos débiles.