

Regresión lineal múltiple 02

Agustin Riquelme y Heriberto Espino

10 de octubre de 2023

Para facilitar el trabajo, se trabajó con funciones.

1. Lectura de la base de datos

A continuación se muestra la lectura de la base de datos.

```
data = Ejemplo_subasta <- read_excel("C:/Users/herie/OneDrive - Fundacion Universidad de las Americas Puebla/Semestre/5 Semestre/Econometria 1/Bases de datos/Ejemplo_subasta.xlsx")
```

2. Construcción del modelo de regresión múltiple con interacciones

A continuación se muestra la construcción del modelo de regresión múltiple con interacciones.

```
mi <- lm( y ~ x1 + x2 + x1 * x2 )
```

3. Extracción de información del modelo

Para la extracción de información del modelo se ocupa la función construida anteriormente con extr.datos(), muestra los valores de los Betas, la desviación estándar, la varianza, y el Test Akaike que muestra que tan buen modelo construimos, mientras menos sea el AIC, mejor modelo es.

```
extr.datos(mi, 0.05, 1)
```

Extracción de datos	
Variable	Valor
β_0	320.4579934
β_1	0.8781425
β_2	-93.2648244
β_3	1.2978458
s	88.9145121
s ²	7905.7904709
AIC	383.7502853

4. Prueba general del modelo

En este fragmento de código, se está llevando la cabo prueba general del modelo,

$$H_0 : \beta_i = 0 \quad v. s. \quad H_1 : \beta_i \neq 0$$

Para la extracción de información del modelo se ocupa la función prueb.glob(), con 5 grados de libertad para conocer la información del modelo.

```
prueb.glob(mi, 0.05)
```

Pruebas global para β_i	
Variable	Valor
FC	193.041095792801
P value F	0.00000000000000000834955968420088
Conclusión	Se rechaza H0, es decir, el modelo es significativo

5. Pruebas individuales

```
prueb.indiv(mi, 0.05, 1)
```

Prueba individual para β_i	
Variable	Valor
P value β_1	0.668961315096404
Concusión	No existe evidencia para afirmar que β_1 sea diferente de cero
P value β_2	0.00416458862648844
Conclusión	Se rechaza H0, es decir, β_2 es significativo para el modelo
P value β_3	0.00000135347405407586
Conclusión	Se rechaza H0, es decir, β_3 es significativo para el modelo

Como en las pruebas individuales, β_1 es mayor que α , tenemos que descartar a X1, y con eso también descartamos a X1*X2, nos quedaríamos con el modelo de X2, que sabemos que es peor que el modelo de regresión lineal multiple que construimos en clases pasadas. Entonces descartamos ambos modelos.

6. Análisis de residuales

```
norm.err(mi, 0.05)
```

Normalidad H0: Err. norm vs H1: Err. no norm	
Test	Valor
Lillie	0.672170094114534
Conclusión	p valor $\geq \alpha$, no rechazamos H0 .
Anderson	0.843546331377835
Conclusión	p valor $\geq \alpha$, no rechazamos H0

7. Pruebas de independencia

```
corr.err(mi, 0.05)
```

Test Durbin-Watson	
Test	Conclusión
DW	2.41898824128436
P valor	0.886181117208715
Conclusión	No hay una correlación en los residuos

8. Construcción del modelo cuadrático simple y completo

A continuación, se están construyendo tres modelos de regresión utilizando el paquete lm():

```
mc1 <- lm( y ~ x1 + I(x1^2) )
mc2 <- lm( y ~ x2 + I(x2^2) )
mcc <- lm( y ~ x1 + x2 + x1*x2 + I(x1^2) + I(x2^2) )
```

El modelo mc1 es un modelo cuadrático simple que predice la variable de respuesta y basándose en la variable predictora x1 y su cuadrado. De manera similar, mc2 es un modelo cuadrático simple que utiliza la variable predictora x2 y su cuadrado para predecir y.

El modelo mcc es un modelo cuadrático completo que incluye las variables x1 y x2, así como sus cuadrados, y también los términos de interacción entre x1 y x2. Este modelo completo permite capturar relaciones más complejas entre las variables predictoras y la variable de respuesta. Estos modelos son útiles para explorar y modelar relaciones no lineales en los datos, proporcionando una comprensión más detallada de la naturaleza subyacente del fenómeno estudiado.

9. Extracción de la información de los modelos

En este fragmento de código, se está realizando la extracción de información clave de los tres modelos de regresión previamente definidos: mc1, mc2 y mcc. La función extr.datos() está siendo utilizada para calcular y mostrar varios valores asociados con cada modelo.

```
extr.datos(mc1, 0.05, 2)
```

Extracción de datos	
Variable	Valor
β_0	-1679.2852498
β_1	31.3191397
β_2	-0.0705421
s	274.7892089
s ²	75509.1093070
AIC	455.0862584

```
extr.datos(mc2, 0.05, 2)
```

Extracción de datos	
Variable	Valor
β_0	-454.89589
β_1	326.10425
β_2	-13.44477
s	357.26414
s ²	127637.66805
AIC	471.88441

```
extr.datos(mcc, 0.05, 3)
```

Extracción de datos	
Variable	Valor
β_0	-331.9260054
β_1	3.2079181
β_2	14.8088791
β_3	-0.0029952
β_4	-4.1790321
β_5	1.1231572
s	86.1018634
s ²	7413.5308737
AIC	383.3215938

10. Prueba general de los modelos

En este fragmento de código, se están llevando a cabo pruebas generales de los modelos de regresión mc1, mc2, y mcc utilizando la función prueb.glob().

La función prueb.glob() realiza una prueba de hipótesis para evaluar la significancia global del modelo de regresión. Esta prueba tiene como objetivo determinar si al menos uno de los coeficientes del modelo es significativamente diferente de cero. En otras palabras, se está probando si el modelo en su conjunto es significativo para predecir la variable de respuesta.

Los argumentos proporcionados a la función son el modelo de regresión y el nivel de significancia, que en este caso es 0.05, indicando un 5% de nivel de significancia.

```
prueb.glob(mc1, 0.05)
```

Pruebas global para β_i	
Variable	Valor
FC	17.2628507318646
P value F	0.0000114298289976184
Conclusión	Se rechaza H0, es decir, el modelo es significativo

```
prueb.glob(mc2, 0.05)
```

Pruebas global para β_i	
Variable	Valor
FC	4.30240203941252
P value F	0.0231069529237575
Conclusión	Se rechaza H0, es decir, el modelo es significativo

```
prueb.glob(mcc, 0.05)
```

Pruebas global para β_i	
Variable	Valor
FC	124.287273520705
P value F	0.00000000000000000269932315943952
Conclusión	Se rechaza H0, es decir, el modelo es significativo

11. Pruebas individuales

```
prueb.indiv(mc1, 0.05, 2)
```

Prueba individual para β_i	
Variable	Valor
P value β_1	0.211778470233371
Concusión	No existe evidencia para afirmar que β_1 sea diferente de cero
P value β_2	0.401244425868745
Conclusión	No existe evidencia para afirmar que β_2 sea diferente de cero

```
prueb.indiv(mc2, 0.05, 2)
```

Prueba individual para β_i	
Variable	Valor
P value β_1	0.0587299547086962
Concusión	No existe evidencia para afirmar que β_1 sea diferente de cero
P value β_2	0.109177985026749
Conclusión	No existe evidencia para afirmar que β_2 sea diferente de cero

```
prueb.indiv(mcc, 0.05, 3)
```

Prueba individual para β_i	
Variable	Valor
P value β_1	0.722828596806087
Conclusión	No existe evidencia para afirmar que β_1 sea diferente de cero
P value β_2	0.813712523907412
Conclusión	No existe evidencia para afirmar que β_2 sea diferente de cero
P value β_3	0.914038657071601
Conclusión	No existe evidencia para afirmar que β_4 sea diferente de cero
P value β_4	0.813712523907412
Conclusión	No existe evidencia para afirmar que β_2 sea diferente de cero
P value β_5	0.914038657071601
Conclusión	Se rechaza H0, es decir, β_3 es significativo para el modelo

Estos resultados sugieren que, en las tres muestras analizadas, la mayoría de los coeficientes β_i no mostraron evidencia suficiente para afirmar que son diferentes de cero. Sin embargo, en el caso de β_3 en la tercera muestra, se encontró evidencia significativa para afirmar que este coeficiente es significativo para el modelo, ya que se rechaza la hipótesis nula. Por lo que quedan descartados los modelos, sin embargo, se mostrará el análisis de residuales y las pruebas de independencia de los residuales.

12. Análisis de residuales

```
norm.err(mc1, 0.05)
```

Normalidad H0: Err. norm vs H1: Err. no norm	
Test	Valor
Lillie	0.956653930771793
Conclusión	p valor $\geq \alpha$, no rechazamos H0 .
Anderson	0.929694155539767
Conclusión	p valor $\geq \alpha$, no rechazamos H0

```
norm.err(mc2, 0.05)
```

Normalidad H0: Err. norm vs H1: Err. no norm	
Test	Valor
Lillie	0.407523130564366
Conclusión	p valor $\geq \alpha$, no rechazamos H0 .
Anderson	0.0849516133701323
Conclusión	p valor $\geq \alpha$, no rechazamos H0

```
norm.err(mcc, 0.05)
```

Normalidad H0: Err. norm vs H1: Err. no norm	
Test	Valor
Lillie	0.680962419684125
Conclusión	p valor $\geq \alpha$, no rechazamos H0 .
Anderson	0.649432450582579
Conclusión	p valor $\geq \alpha$, no rechazamos H0

13. Prueba de independencia

```
corr.err(mc1, 0.05)
```

Test Durbin-Watson	
Test	Conclusión
DW	1.85698910400399
P valor	0.341075857488755
Conclusión	No hay una correlación en los residuos

```
corr.err(mc2, 0.05)
```

Test Durbin-Watson	
Test	Conclusión
DW	2.06042043291706
P valor	0.585323280697026
Conclusión	No hay una correlación en los residuos

```
corr.err(mcc, 0.05)
```

Test Durbin-Watson	
Test	Conclusión
DW	2.42070406080497
P valor	0.903500506039086
Conclusión	No hay una correlación en los residuos