

FACTORES DETERMINANTES EN EL CONSUMO DE TABACO ENTRE LOS JÓVENES



Marina Pérez Bujalance



ECONOMETRÍA II

ÍNDICE

1. Resumen
2. Motivación
3. Introducción
4. Datos
 - 4.1 Descripción de los datos
 - 4.2 Estadísticos descriptivos
 - 4.3 Matriz de correlaciones
 - 4.4 Gráfico de dos variables
5. Resultados
 - 5.1 Especificación del modelo
 - A. Modelo lineal de probabilidad (MLP)
 - B. Logit
 - 5.2 Contraste de las hipótesis básicas
 - A. Normalidad
 - B. Multicolinealidad
 - C. Heterocedasticidad
 - D. Autocorrelación
6. Discusión
7. Conclusiones
8. Bibliografía

1. RESUMEN:

El consumo de tabaco por parte de los jóvenes es un tema de gran interés tanto para padres como para Gobiernos ya que constituye una importante fuente de muertes. Por ello este trabajo tiene por objetivo el estudio de las variables dominantes en el consumo de tabaco entre estudiantes de Economía de la Universidad de Granada. Intenta entender cuáles son las variables explicativas de este fenómeno atendiendo a factores como la edad, los ingresos o si las personas cercanas al individuo fuman.

Conocer estas variables nos permitirá saber qué planes de contingencia serán eficaces para bajar el número de fumadores pero ¿Por qué el Estado está interesado en este temas si el tabaco está sujeto a altos tipos impositivos? Lo cierto es que los estudios revelan que los costes son superiores a los ingresos lo que quiere decir que eliminando fumadores el Estado está disminuyendo costes como pueden ser los sanitarios.

2. MOTIVACIONES:

Según la OMC el tabaco mata cada año a 7 millones de personas, incluyendo a fumadores activos y pasivos. Ante este fenómeno que perdura en el tiempo, instituciones internacionales como la OMS han decidido tomar cartas en el asunto, entre sus actividades destaca el Día Mundial Sin Tabaco donde reúne delegados de países europeos para poner de manifiesto los riesgos para la salud del tabaquismo y defender políticas que reduzcan el consumo de tabaco.

En España existe en conocido Comité Nacional para la Prevención del Tabaquismo (CNT) cuyos objetivos son similares, la prevención y control del tabaquismo.

En nuestro país como en el resto de países europeos se llevan a cabo políticas para la reducción del consumo. No obstante debemos ser conscientes de la presencia de numerosos lobbies tabacaleros y puertas. Volviendo al tema de las políticas, destaca la Ley antitabaco de 2011 con la que se prohibió el consumo de tabaco en espacio de usos colectivos como por ejemplo los bares. Sin embargo los más jóvenes siguen iniciándose en el consumo de tabaco y aquí es donde empieza este trabajo cuyo objetivo es localizar las variables que influyen en el consumo de tabaco por parte de los más jóvenes. Pese a que se intenta concienciar de sus efectos negativos mediante: anuncios publicitarios, charlas en institutos o envoltorios que muestran los efectos dañinos, muchos adolescentes deciden

probar la sustancia. Parte de la culpa de esto puede recaer en las redes sociales, en las cuales los jóvenes compartimos gran parte de nuestra vida. En ellas encontramos numerosas fotos donde adolescentes exhiben su consumo sin ningún tipo de pudor y sin recibir apenas críticas por parte de los visitantes de esta plataforma. Esto me lleva a pensar que en este tipo de redes el consumo de tabaco no es un tema tabú como muchos considerábamos que era, al contrario es un símbolo socialmente aceptado.

Hablando en términos económicos el tabaquismo supone un gran coste para el estado. Lo que solemos escuchar es que al gobierno le interesa mantener el consumo de tabaco porque su carga en impuestos es mayor que su coste sanitario pero esto no es así.

Por el lado de la carga fiscal concluimos que de cada 1€ gastado: 0.80€ van para el estado en concepto de IVA y de impuesto especial

Por el lado de los costes señalamos principalmente el coste sanitario, bastante elevado porque incluyen medicación, diagnósticos, tratamientos...

En resumen, la razón porque la que existen numerosos estudios sobre las variables explicativas del fenómeno es porque el estado a parte de querer una población sana, también la quiere productiva y sin costes derivados del tabaco. Luego partiendo de los resultados el estado realiza campañas preventivas con las que se quiere reducir el consumo de tabaco. A la larga estas campañas no son vistas como un coste sino como una inversión: este dinero retornará al estado con menos gente enferma y por tanto con menos costes que sufragar.

Dicho todo esto no debemos olvidar que estamos hablando de vidas humanas por lo que su valor no solo se mide en términos capital, pero sí es interesante estudiar su impacto en la economía.

3. INTRODUCCIÓN:

En este estudio he descrito una serie de variables que se piensa que pueden influir en el consumo de tabaco entre los universitarios con el propósito de obtener conclusiones que nos ayuden a entender un poco mejor que lleva a los estudiantes a fumar y a estabilizar su consumo.

Para realizar el modelo hemos tomado como ayuda algunos artículos, destacan:

- Artículo (Cardenal, C. A., & i Adell, M. N. ,1995) elaborado en Cataluña sobre 1816 individuos donde apunta al ambiente familiar y de amigos como variable explicativa,

es decir, tener padres fumadores sería un factor determinante en el inicio al consumo y tener amigos fumadores es clave para mantener el hábito.

- Artículo (Puente, D., Zabaleta-del-Olmo, E., Pueyo, M. J., Saltó, E., Marsal, J. R., & Bolívar, B. ,2013) también elaborado en Cataluña en 2012 con 9340 observaciones, afirma que aumenta la probabilidad de fumar: tener bajo rendimiento escolar, tener problemas de ánimo, consumir alcohol o existencia de familiares fumadores entre otros
- En este estudio(Nuño-Gutiérrez, B. L., Alvarez-Nemegyei, J., León, E. M. D., & Rasmussen-Cruz, B. ,2005). no se concretan diferencias en el patrón de consumo entre hombres y mujeres pero sí distingue entre consumo de tabaco materno y paterno, siendo este último influyente en la probabilidad de que los jóvenes fumen, atribuye este resultado al papel de autoridad que este representa en el núcleo familiar.
- Leyendo el estudio (Inglés, C. J., Delgado, B., Bautista, R., Torregrosa, M. S., Espada, J. P., García-Fernández, J. M., ... & García-López, L. J. 2007)realizado para la red de revistas científicas de América Latina, sus odd ratios afirman que la probabilidad de ser fumador es mayor para aquellos adolescentes que presentan un comportamiento antisocial o tienen inestabilidades emocionales que para los que no lo presentan.
- En otros estudios(Nebot, M., Tomás, Z., Ariza, C., Valmayor, S., López, M. J., & Juárez, O. 2004) separa los factores que incitan a fumar a lo chicos de aquellos que incitan a fumar a las chicas. Entre ellos los principales factores son: la intención de fumar en el futuro, baja auto eficacia para resistir presiones hacia el consumo, el consumo de tabaco de los hermanos o pasar tiempo en bares. Mientras que para ellas destaca tener baja autoestima, el tabaquismo de los hermanos o pasar parte de su tiempo libre en discotecas.

4. DATOS:

4.1 Descripción de los datos

Los datos de este trabajo han sido recogidos por una encuesta online realizada en el mes de marzo en la que han participado 87 individuos pertenecientes a los diferentes cursos del grado GECO en la UGR. Los datos son de corte transversal y han sido tratados con el programa de Gretl.

VARIABLE	TIPO	DESCRIPCIÓN	UNIDAD MEDIDA	SIGNO ESPERADO
Fumador	V.explicada Binaria	1-Fumador 0- No fumador		
Edad	V.explicativa Continua		años	Positivo Aumenta la edad, aumenta la probabilidad de ser fumador
Sexo	V.explicativa Binaria	1- Hombre 0-Mujer		Positivo Los hombres tiene una mayor probabilidad de ser fumadores
Dinero destinado al ocio	V.explicativa Continua	Representa la cantidad de dinero semanal que el individuo dedica al ocio	euros	Positivo Aumento en la cantidad de dinero, aumento en la probabilidad
Padres fumadores	V.explicativa Binaria	1- Ambos progenitores o alguno de ellos fuman 0- Ningún progenitor fuma		Positivo Si los progenitores fuman la probabilidad de que fumen sus hijos aumentará
Amigos fumadores	V.explicativa Binaria	1- Al menos uno de sus amigos fuma 0- Ninguno de sus amigos fuma		Positivo Si sus amigos fuman, la probabilidad de fumar del individuo aumentará
Presión social	V.explicativa Binaria	1- Siente presión social(al menos una vez alguien le ha incitado a fumar) 0- No siente presión social		Positivo Si siente presión por fumar, aumentará su probabilidad
Salud	V.explicativa Binaria	1- conoce los daños del tabaquismo 0- no conoce los efectos		Negativo Si conoce los efectos perjudiciales su probabilidad disminuirá
Distancia	V.explicativa Binaria	1- cercanía a puntos de venta (menos de 5 minutos)		Positivo Si tiene cerca puntos de venta aumentarla la probabilidad de

		0- Lejanía a puntos de venta(más de cinco minutos)		que sea fumador
Redes sociales	V.explicativa Binaria	1- Ve escenas de fumadores en sus redes sociales al menos una vez a la semana) 0- No ve este tipo de escenas		Positivo Cuanto más escenas de fumadores vea, más aumentará su probabilidad
Estrés	V.explicativa Binaria	1- Sufre estrés a menudo (dolores de cabeza, ansiedad...) 0- no sufre estrés		Positivo Si sufre estrés a frecuentemente su probabilidad de ser fumador aumenta

4.2 Estadísticos descriptivos

Estadísticos principales, usando las observaciones 1 - 87

Variable	Media	Mediana	D. T.	Mín	Máx
Sexo	0.414	0.00	0.495	0.00	1.00
Edad	21.4	21.0	3.40	18.0	40.0
Dinero	20.8	15.0	17.4	0.00	100.
Padres	0.379	0.00	0.488	0.00	1.00
Amigos	0.690	1.00	0.465	0.00	1.00
Presionsocial	0.126	0.00	0.334	0.00	1.00
Salud	0.943	1.00	0.234	0.00	1.00
Redessociales	0.724	1.00	0.924	0.00	8.00
Distancia	0.471	0.00	0.502	0.00	1.00
EstrAs	0.747	1.00	0.437	0.00	1.00

Estos datos serán de interés para calcular la probabilidad de que ocurra el suceso para el individuo representativo de este estudio.

4.3 Matriz de correlaciones

Matriz de correlación =

	Sexo	Edad	Dinero	Padres	Amigos	Presión social	Salud	Redes sociales	Distancia	Estrés	
Sexo	1	-0.11	0.15	-0.07	0.008	-0.10	0.006	0.22	0.09	0.005	Sexo
Edad		1	0.01	-0.11	-0.07	-0.01	-0.003	0.02	-0.28	0.03	Edad
Dinero			1	0.03	0.06	-0.02	-0.01	0.06	0.02	0.12	Dinero
Padres				1	0.06	-0.01	-0.01	-0.12	0.06	0.01	Padres
Amigos					1	-0.11	-0.16	-0.03	0.08	0.06	Amigos
Presión social						1	-0.20	0.07	-0.08	0.6	Presión social
Salud							1	0.03	-0.06	0.08	Salud
Redes sociales								1	-0.11	0.11	Redes sociales
Distancia									1	0.01	Distancia
Estrés										1	Estrés

```

gret: matriz de correlación
Coeficientes de correlación, usando las observaciones 1 - 87
valor crítico al 5% (a dos colas) = 0.2108 para n = 87

Sexo          Edad          Fumas          Dinero
1.0000        -0.1161        0.1018        0.1576
              1.0000        0.0403        0.0135
              1.0000        1.0000        0.2087
              1.0000        1.0000        1.0000
              Sexo

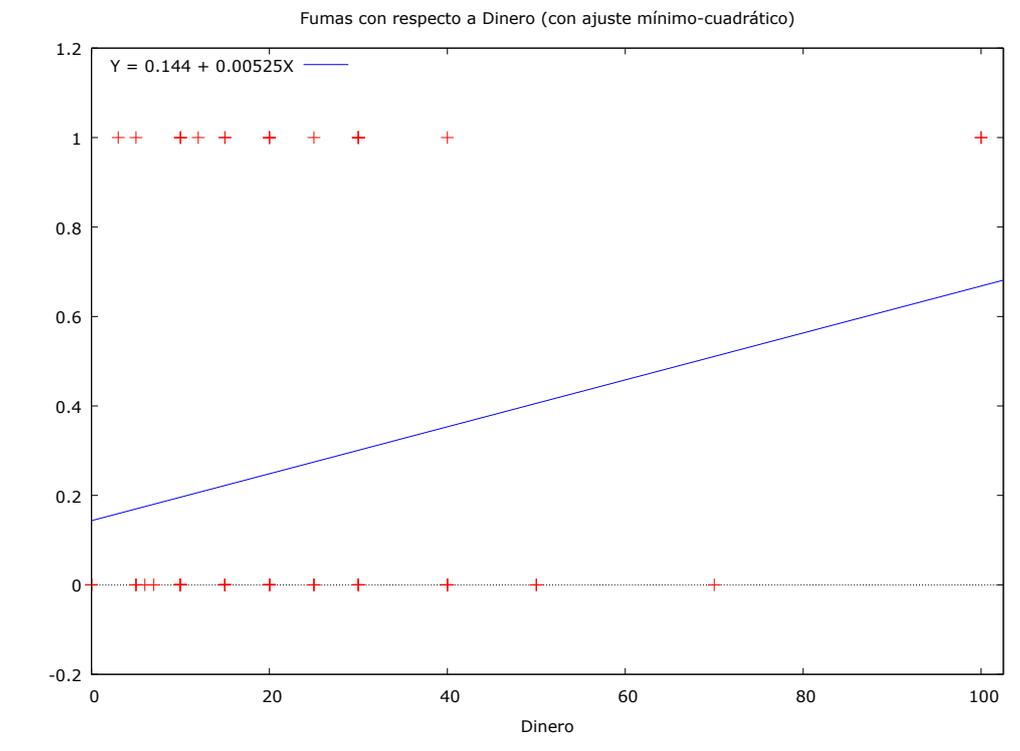
Padres        Amigos        Presionsocial  Salud
-0.0796      0.0087        -0.1090      0.0069
-0.1103      -0.0762      -0.0196      -0.0032
-0.0188      0.2188        0.0174        0.0300
0.0306       0.0627        -0.0285      -0.0166
1.0000       0.0636        -0.0123      -0.0105
              1.0000        -0.1186      -0.1656
              1.0000        1.0000      -0.2033
              1.0000        1.0000      1.0000
              Presionsocial
              Salud

Redessociales  Distancia  EstrAs
0.2270         0.0951     0.0056
0.0279        -0.2832     0.0300
0.0308         0.1924     0.0951
0.0630         0.0208     0.1230
-0.1263       0.0687     0.0188
-0.0392       0.0858     0.0670
0.0766        -0.0820     0.0622
0.0334        -0.0637     0.0836
1.0000        -0.1176     0.1132
              1.0000     0.0195
              1.0000     1.0000
              EstrAs
  
```

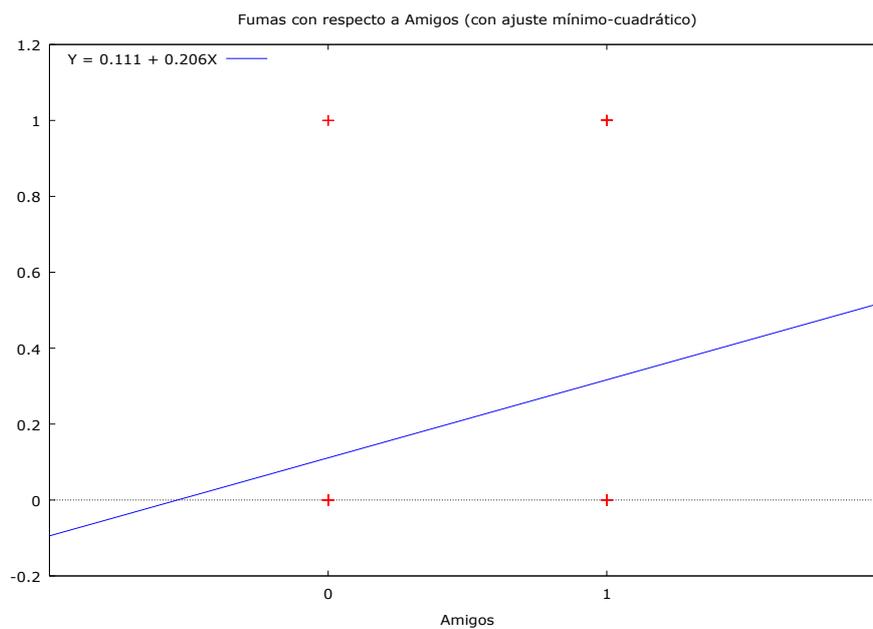
Podemos concluir que no existe correlación entre las variables explicativas puesto que no existen valores altos próximos a 0.8, si encontrásemos alguno estaríamos frente a un problema de multicolinealidad

4.4. Gráfico de dos variables

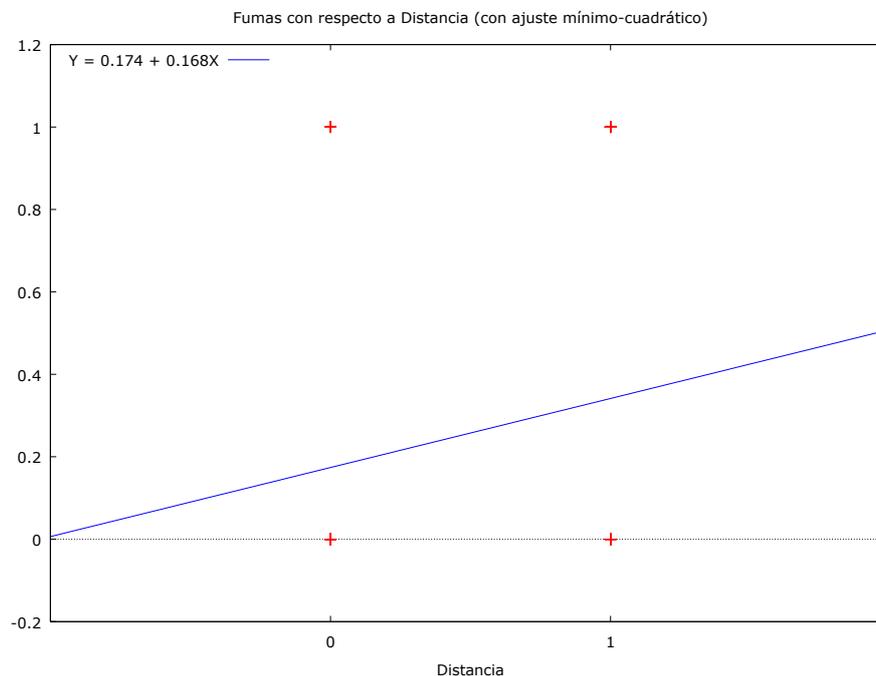
- Fumar y los ingresos



- Fumar y la existencia de amigos fumadores



- Fumar y la distancia a los puntos de venta del tabaco



Estas son las únicas gráficas donde Gretl establece una relación. Todas ellas muestran un relación directa si aumenta la cantidad de dinero o las variables amigos y distancia toman el valor 1 la probabilidad aumentar

5. RESULTADOS

5.1 Especificación del modelo

Estimaremos el modelo por el Modelo Lineal de Probabilidad(MLP) y por Logit. La estimación por MLP es intuitiva ya que sus parámetros muestran la propensión marginal pero tiene varios inconvenientes entre ellos destacan:

En primer lugar que la probabilidad puede que no se encuentre en el intervalo (0,1). En segundo lugar las perturbaciones no cumplen el supuesto de normalidad, luego se invalida el uso de estadísticos para el contraste de hipótesis porque estos se basan en dicho supuesto. En tercer lugar las perturbaciones son heterocedásticas, es decir, su varianza depende de las observaciones.

A. MODELO LINEAL DE PROBABILIDAD

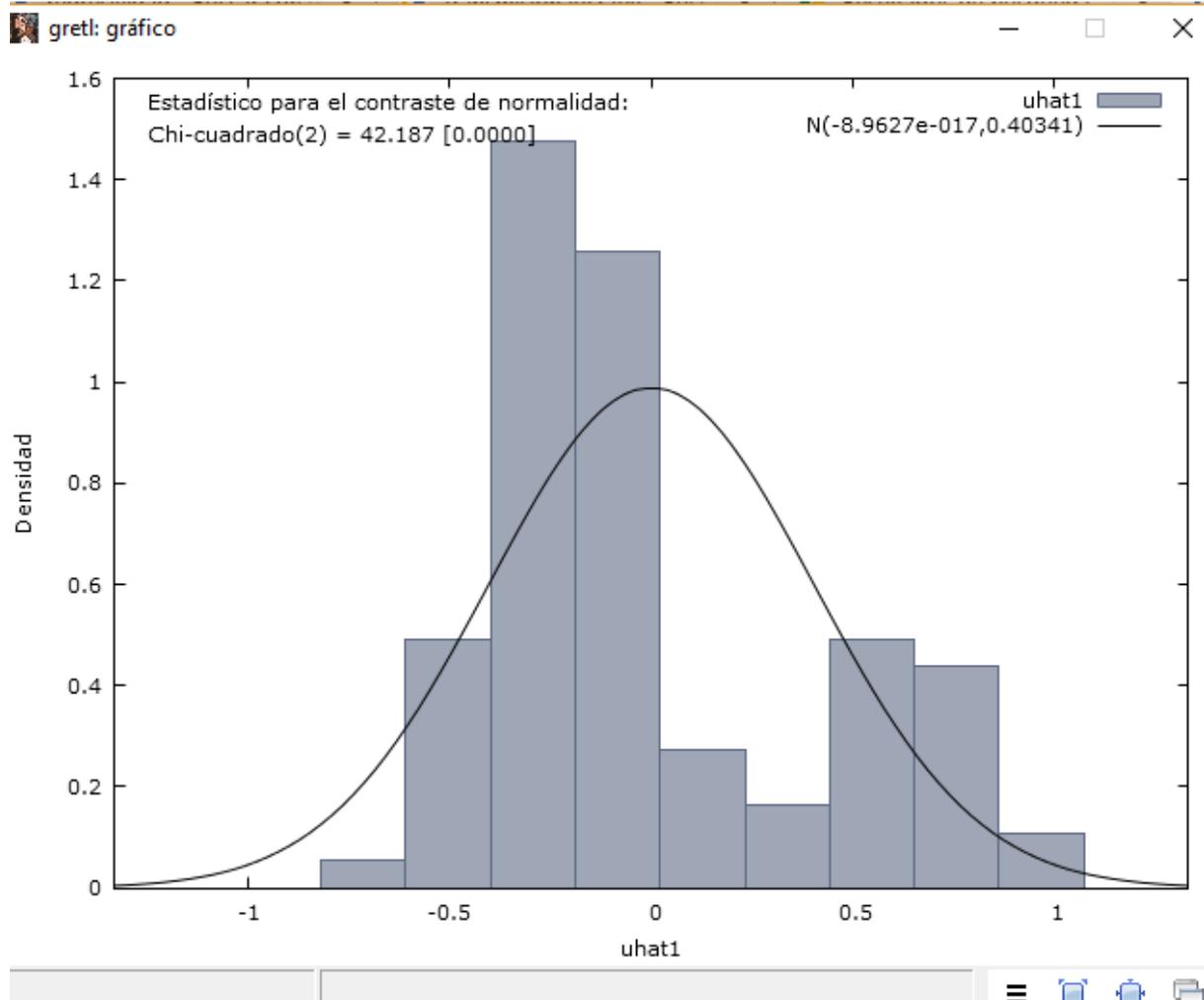
Modelo 1: MCO, usando las observaciones 1-87
Variable dependiente: Fumas

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>Estadístico t</i>	<i>valor p</i>	
const	-0.642177	0.420825	-1.526	0.1312	
Sexo	0.0605383	0.0993391	0.6094	0.5441	
Edad	0.0155503	0.0144089	1.079	0.2839	
Dinero	0.00446340	0.00272373	1.639	0.1054	
Padres	-0.0276540	0.0966760	-0.2860	0.7756	
Amigos	0.211951	0.103311	2.052	0.0437	**
Presionsocial	0.121584	0.146213	0.8316	0.4083	
Salud	0.184978	0.208209	0.8884	0.3771	
Redessociales	0.00778077	0.0528901	0.1471	0.8834	
Distancia	0.186724	0.0979988	1.905	0.0605	*
EstrAs	0.0346053	0.108717	0.3183	0.7511	
Media de la vble. dep.	0.252874	D.T. de la vble. dep.		0.437179	
Suma de cuad. residuos	13.99540	D.T. de la regresión		0.429127	
R-cuadrado	0.148532	R-cuadrado corregido		0.036496	
F(10, 76)	1.325758	Valor p (de F)		0.232365	
Log-verosimilitud	-43.96534	Criterio de Akaike		109.9307	
Criterio de Schwarz	137.0557	Crit. de Hannan-Quinn		120.8531	

Bajo el MLP:

- Obtenemos como significativas las variables: Amigos al 5% de significación (si el individuo tiene amigos fumadores la probabilidad de que fume es un 21% mayor) y distancia al 10% de significación (si el individuo tiene normalmente puntos de venta cercanos la probabilidad de que fume aumenta en un 18%)
- El $R^2 = 0.14$ su valor es bajo, el modelo no es demasiado bueno
- Estudiar los resultados de los test no tiene mucha credibilidad porque no hay normalidad en las perturbaciones, condición necesaria para la fiabilidad de los mismos

Para ver que no hay normalidad en las perturbaciones podemos ver el siguiente gráfico y usar el estadístico Jarque - Bera



gretl: contraste de normalidad

Contraste de normalidad de uhat1:

Contraste de Doornik-Hansen = 42.1866, con valor p 6.90699e-010

W de Shapiro-Wilk = 0.86448, con valor p 2.07823e-007

Contraste de Lilliefors = 0.22177, con valor p ≈ 0

Contraste de Jarque-Bera = 11.9201, con valor p 0.00257976

Como el p-valor del contraste es menor de 0.01 podemos concluir que rechazamos la hipótesis nula de normalidad.

B. LOGIT

Estimaremos ahora el modelo por el método Logit donde la probabilidad si se encuentra en el intervalo (0,1) se trata de un modelo de regresión no lineal donde solo podremos interpretar el signo de los parámetros. El resultado es el siguiente:

Modelo 2: Logit, usando las observaciones 1-87
Variable dependiente: Fumas
Desviaciones típicas basadas en el Hessiano

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>valor p</i>	
const	-6.88652	2.60464	-2.644	0.0082	***
Sexo	0.445495	0.579020	0.7694	0.4417	
Edad	0.101877	0.0803444	1.268	0.2048	
Dinero	0.0243592	0.0147455	1.652	0.0985	*
Padres	-0.199206	0.570166	-0.3494	0.7268	
Amigos	1.45422	0.714411	2.036	0.0418	**
Presionsocial	0.784526	0.862699	0.9094	0.3631	
Salud	1.02793	1.23130	0.8348	0.4038	
Redessociales	0.0333614	0.291749	0.1143	0.9090	
Distancia	1.17425	0.596587	1.968	0.0490	**
EstrAs	0.165340	0.678548	0.2437	0.8075	
Media de la vble. dep.	0.252874	D.T. de la vble. dep.	0.437179		
R-cuadrado de McFadden	0.142226	R-cuadrado corregido	-0.081370		
Log-verosimilitud	-42.19895	Criterio de Akaike	106.3979		
Criterio de Schwarz	133.5229	Crit. de Hannan-Quinn	117.3203		

Número de casos 'correctamente predichos' = 71 (81.6%)

f(beta'x) en la media de las variables independientes = 0.437

Contraste de razón de verosimilitudes: Chi-cuadrado(10) = 13.9939 [0.1733]

La especificación del modelo será:

$Z_i = -6.88 + 0.44SEXO + 0.10EDAD + 0.024DINERO -$

$0.19PADRES + 1.45AMIGOS + 0.784PSOCIAL + 1.02SALUD + 0.033REDES + 1.17DISTANCIA + 0.165ESTRÉS.$

Obtenemos como **variables significativas**:

- Término independiente (al 1% de significación). Que sea o no significativo no tiene mucha relevancia.
- Dinero (al 10% de significación) con un signo positivo: un aumento en la cantidad de dinero semanal para el ocio se traducirá en un aumento en la probabilidad de ser fumador

- Amigos (al 5% de significación) con signo positivo: si el individuo se rodea de amigos fumadores la probabilidad de que él lo sea aumenta
- Distancia (al 5% de significación) con signo positivo: si el individuo tiene puntos de venta cercanos la probabilidad de que sea fumador aumenta

Los coeficientes de las variables significativas tienen el signo esperado. Destacan dos variables con signos distintos al esperado: padres y salud.

En el modelo Logit podemos hallar:

- Los **Odds**, un coeficiente que se interpreta como el número de veces que es más probable que ocurra el fenómeno o suceso de interés ($Y=1$ que sea fumador) frente a que no ocurra ($Y=0$ que no sea fumador). Para calcularlos necesitamos las características de un individuo, por tanto podríamos usar el individuo medio pero como casi todas nuestras variables explicativas son binarias este individuo no tiene sentido porque si el valor 1 significa tener padres fumadores y el 0 no tenerlos, el 0.3 del individuo medio no indicaría nada. ¿Qué podemos hacer? Podemos usar el valor medio de nuestras variables continuas (edad y dinero) y calcular la moda para las variables binarias. Estos datos ya los hemos calculado previamente en el apartado 4.2 de Estadísticos descriptivos.

Por tanto calculamos el odd para un individuo cuyo sexo es mujer, su edad es de 21.4 años, dedica semanalmente al ocio 20.8€, sus padres no fuman, sus amigos si fuman, no siente presión social para fumar, conoce los efectos a largo plazo en la salud de ser fumador, considera que ve al menos una vez a la semana escenas de fumadores en las redes sociales, tiene puntos de venta de tabaco a más de cinco minutos y sufre síntomas de estrés.

$$Z_i = -6.88 + 0.10 \cdot 21.4 + 0.024 \cdot 20.8 + 1.45 \cdot 1 + 1.02 \cdot 1 + 0.033 \cdot 1 + 0.165 \cdot 1 = -1.5728$$

$$\text{Odd} = e^{Z_i} = e^{-1.5728} = 0.2074$$

Quiere decir que el individuo medio de nuestro estudio tiene 0.2074 veces más de posibilidades de fumar que de no hacerlo

-Los **efectos marginales**: teniendo las características es este individuo representativo de nuestro estudio podremos calcular el efecto marginal de algunas variables

<u>E.marginal del sexo(binaria)</u>	Prob. cuando es mujer y todo lo demás constante Zi=-1.5728 P(Ym)=0.1718	EM= P(Yh)-P(Ym)= 0.2436-0.118=0.0718 Es un 7.18% más probable que fume un hombre que un mujer dadas las características del individuo representativo de nuestro estudio
	Prob. cuando es hombre y todo lo demás constante Zi=-1.1328 P(Yh)=0.2436	
E. Marginal del dinero(continua)	EM= $\beta \cdot P(Y)(1-P(Y)) = 0.02435 \cdot 0.1718 \cdot 0.8281 = 0.00346$ El aumento en una unidad monetaria en la cantidad de dinero semanal dedicado al ocio aumentará para prob. de fumar en un 0.34% la probabilidad de fumar	

- los **Odds ratio** que se interpretan como cuantas veces es más probable que ocurra que el individuo sea fumador frente a no serlo cuando la variable explicativa aumenta en una unidad manteniendo los demás constante. Por ejemplo:

ODD RATIO de dinero = $e^{0.024} = 1.0242$ Un individuo que aumenta en una unidad el dinero semanal destinado al ocio es 1.0242 veces más probable que fume a que no lo haga que otro individuo.

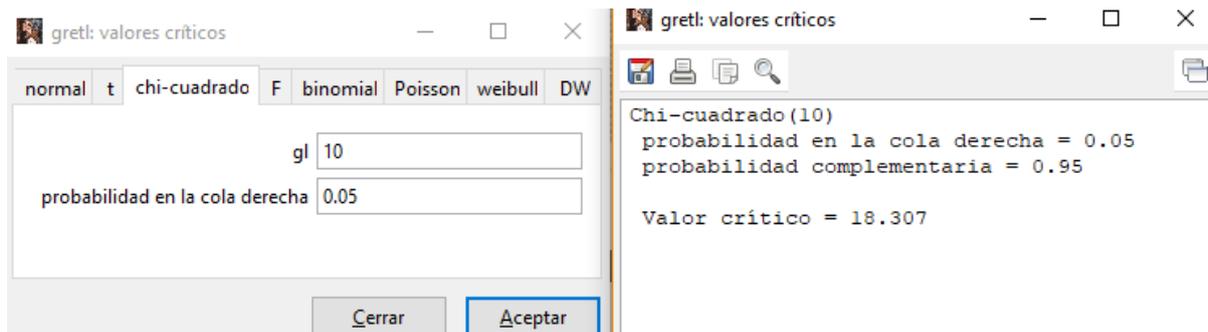
ODD RATIO de amigos= $e^{1.454} = 4.280$ Un individuo que tenga amigos fumadores frente a otros que no los tenga incrementará en 328.02% $[(4.208-1) \cdot 100]$ las posibilidades a favor de ser fumador

ODD RATIO de distancia = $e^{1.17} = 3.2219$ Es 3.22 veces más probable que fume un individuo que tiene puntos de venta cerca frente a otro que no los tenga.

En cuanto a la **significación conjunta del modelo** usamos el contraste de razón de verosimilitudes:

$$\chi_{\text{exp}}^2 = -2 \ln \left(\frac{L_r}{L} \right) \sim \chi_q^2$$

Su valor experimental en la salida de Gretl es 13.99 y el valor teórico es 18.307



Encontramos que el valor teórico es mayor que el valor experimental por lo que no rechazamos la hipótesis nula, es decir el modelo no es significativo en su conjunto.

Esta idea coincide con el resultado de la R^2 de McFadden (0.14) próximo a cero lo que quiere decir que el modelo global es muy similar al modelos restringido es decir las variables no mejoran el modelo, esto quiere decir que este modelo **no** tiene capacidad explicativa

No obstante, llama la atención la buena capacidad de predicción del modelo que obtiene un número de casos correctamente predichos de 71 de un total de 87.

Podemos hacer un contraste sobre el modelo Logit inicial donde las variables no significativas tomen el valor 0 y obtenemos como resultado:

Modelo 3: Logit, usando las observaciones 1-87
Variable dependiente: Fumas
Desviaciones típicas basadas en el Hessiano

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>valor p</i>	
const	-3.06007	0.783465	-3.906	<0.0001	***
Dinero	0.0246245	0.0140773	1.749	0.0803	*
Amigos	1.23877	0.687595	1.802	0.0716	*
Distancia	0.912408	0.537023	1.699	0.0893	*
Media de la vble. dep.	0.252874	D.T. de la vble. dep.		0.437179	
R-cuadrado de McFadden	0.108040	R-cuadrado corregido		0.026732	
Log-verosimilitud	-43.88080	Criterio de Akaike		95.76159	
Criterio de Schwarz	105.6252	Crit. de Hannan-Quinn		99.73337	

Número de casos 'correctamente predichos' = 65 (74.7%)
f(beta'x) en la media de las variables independientes = 0.437
Contraste de razón de verosimilitudes: Chi-cuadrado(3) = 10.6302 [0.0139]

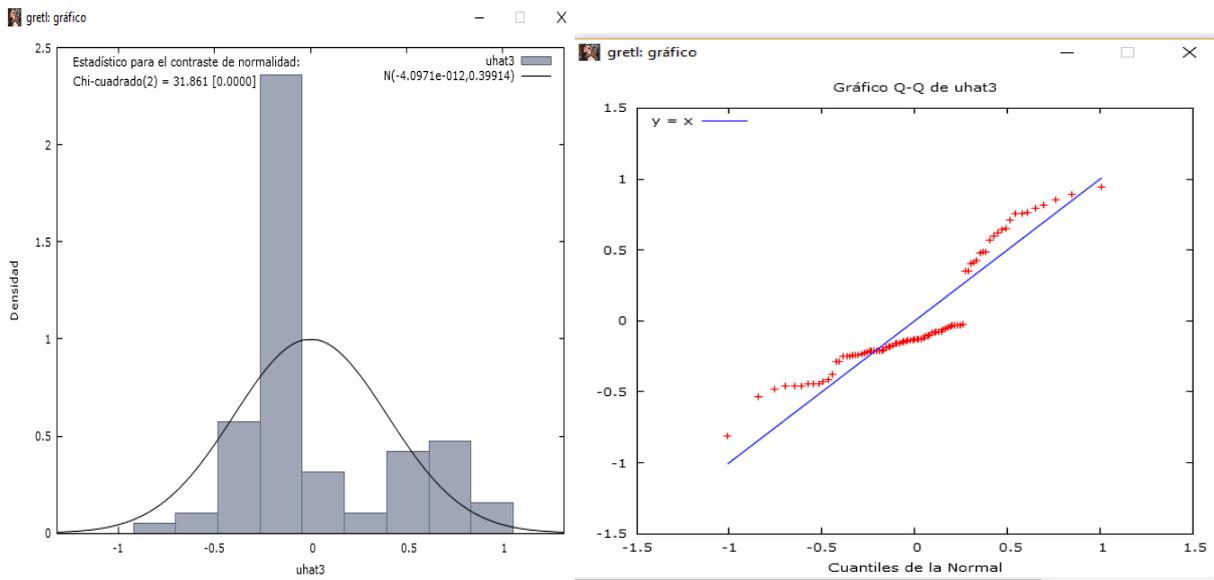
Pese a que el valor de la R^2 de McFadden disminuye pasando del 0.14 al 0.10 nuestro modelo se vuelve significativo en su conjunto con una p-valor del contraste de

verosimilitudes de $0.0139 < 0.05$. Además por el criterio de Akaike (que mide la calidad del modelo) preferimos este último modelo puesto que el valor de este criterio disminuye.

5.2 Contrastes de las hipótesis básicas:

A. NORMALIDAD DE LAS PERTURBACIONES

El gráfico Q-Q no permite analizar si las perturbaciones siguen una distribución normal de manera visual



Si las perturbaciones siguieran una distribución normal la nube de puntos debería ajustarse perfectamente a la bisectriz azul que muestra la gráfica. Si queremos obtener datos de manera más exacta y empírica podemos usar el estadístico Jarque - Bera

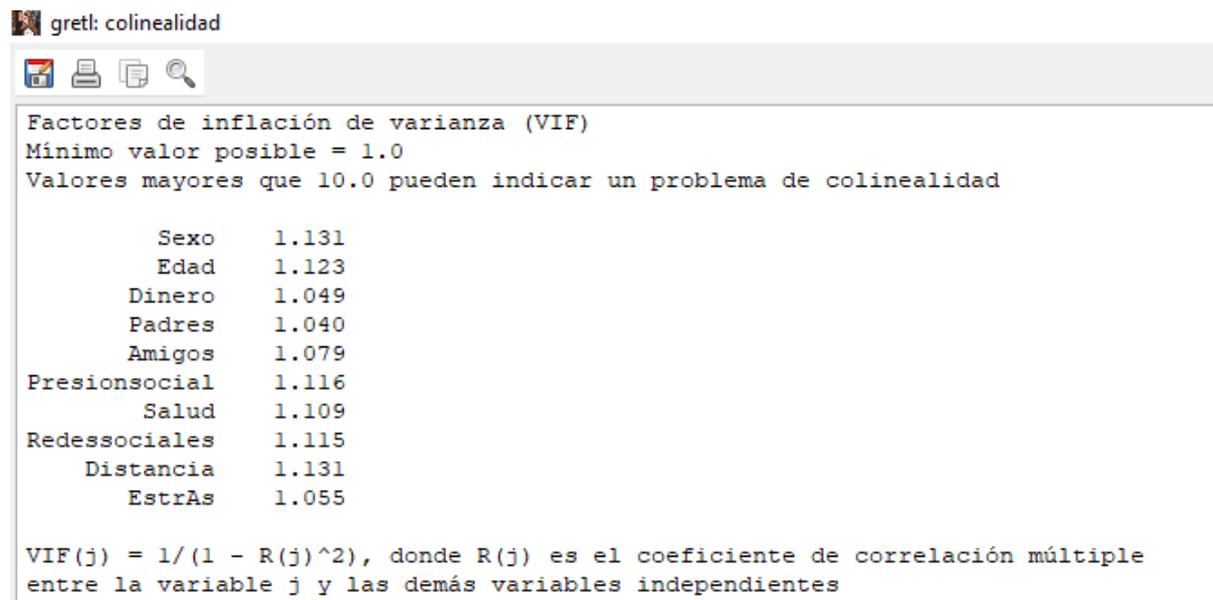
```
gretl: contraste de normalidad
Contraste de normalidad de uhat3:
Contraste de Doornik-Hansen = 31.861, con valor p 1.20636e-007
W de Shapiro-Wilk = 0.863597, con valor p 1.92785e-007
Contraste de Lilliefors = 0.26853, con valor p ~ = 0
Contraste de Jarque-Bera = 11.5812, con valor p 0.0030561
```

El p-valor es 0.003 por lo que con un nivel de significación del 1% podemos rechazar la hipótesis nula de normalidad, concluimos que las perturbaciones no siguen una distribución normal

Obtenemos el mismo resultado con el gráfico que de manera empírica

B. MULTICOLINEALIDAD:

Ya hemos hecho la matriz de correlación que nos indicó que no existía relación lineal entre las variables pero podemos calcular también Factor de Inflación de la Varianza (FIV)



```
gretl: colinealidad
Factores de inflación de varianza (VIF)
Mínimo valor posible = 1.0
Valores mayores que 10.0 pueden indicar un problema de colinealidad

      Sexo      1.131
      Edad      1.123
      Dinero     1.049
      Padres     1.040
      Amigos     1.079
Presionsocial  1.116
      Salud     1.109
Redessociales 1.115
      Distancia 1.131
      EstrAs    1.055

VIF(j) = 1/(1 - R(j)^2), donde R(j) es el coeficiente de correlación múltiple
entre la variable j y las demás variables independientes
```

Vemos que no existen valores altos.

6. DISCUSIÓN

Comparando los resultados obtenidos por nuestro modelo logit con otros estudios realizados, vemos que nuestro modelo solo coincide con los artículos leídos en señalar a la variable amigos como significativa mientras que un factor como el estrés que sí aumenta el riesgo de fumar semanalmente según el estudio Cano, A. J., Solanas, S. E., Marí-Klose, M., & Marí-Klose, P. (2012), en nuestro modelo no es considerado como un factor explicativo.

Este hecho no debe llamarnos la atención puesto que somos conscientes de todas las limitaciones de nuestro modelo. Por un lado un coeficiente de correlación bajo lo que

quiere decir, que nuestro modelo apenas tiene capacidad para explicar si un determinado individuo con ciertas características fumará o no.

Por otro lado la falta de normalidad en la perturbaciones que invalida los resultados de nuestros test t, es decir que estamos afirmando que una variable es explicativa con un test que con cumple uno de los principios en los que se sustenta. Y por último la falta de variables explicativas, para explicar un fenómeno como este debemos de tener en cuenta muchas otras variables como pueden ser: los resultados académicos, si vive en un zona de alto o bajo poder adquisitivo, el estado de ánimo...

Gran parte de la culpa de que nuestro modelo esté tan limitado recae en que tenemos una muestra sesgada, no podemos pretender encontrar un modelo con alto nivel explicativo sobre una población en este caso universitarios del Grado en Economía de la Universidad de Granada solo con 87 observaciones. A esto debemos sumarle que el método de obtención de datos (una encuesta online) favorece el falseamiento de las respuestas.

Por lo tanto, este estudio nos ayuda a comprobar que cuando se hace un estudio con una muestra escasa y poco representativa de la población, los resultados obtenidos no son muy fiables e incluso de nulo valor porque nuestro modelo de regresión pese a tener algunas variables significativas, en su conjunto no es capaz de predecir el fenómeno que nos concierne.

Luego las principales recomendaciones para este modelo serían ampliar el número de observaciones, aumentar el número de variables explicativas a tener en cuenta, recolectar datos de manera más fiable...

7. CONCLUSIONES

Después de tratar nuestros datos con Gretl podemos concluir en que las variables dinero, distancia a puntos de venta, y amigos son variables explicativas de nuestra variable explicada (la probabilidad de ser fumador). Destacando la variable amigos y conectado con nuestra motivación, mantener un entorno de amigos sanos disminuirá nuestras probabilidades de fumar. Es decir, las campañas preventivas promulgadas por el gobierno puede que solo capten a ciertos jóvenes pero el resto de jóvenes que no son captados por la campaña, pueden que sí lo sean por sus amigos que dejaron de fumar por la acción del gobierno. Lo que quiero decir es que el gasto que lleve a cabo el gobierno puede ser un efecto expansivo si variables como amigos o familia son realmente significativas.

Lamentablemente con los resultados de nuestro modelo no nos encontramos en un escenario muy deseable ya que aunque nuestro modelo es significativo en su conjunto, su

R² de McFadden es muy próxima a 0. Esto se debe entre otras a las limitaciones citadas previamente en la discusión y que debemos mejorar.

8. BIBLIOGRAFÍA

1. i Cardenal, C. A., & i Adell, M. N. (1995). Factores asociados al consumo de tabaco en una muestra de escolares de enseñanza primaria y secundaria. *Gaceta Sanitaria*, 9(47), 101-109.
2. Puente, D., Zabaleta-del-Olmo, E., Pueyo, M. J., Saltó, E., Marsal, J. R., & Bolívar, B. (2013). Prevalencia y factores asociados al consumo de tabaco en alumnos de enseñanza secundaria de Cataluña. *Atención Primaria*, 45(6), 315-323.
3. Nuño-Gutiérrez, B. L., Alvarez-Nemegyei, J., León, E. M. D., & Rasmussen-Cruz, B. (2005). Prevalencia y factores asociados al consumo de tabaco en adolescentes de una preparatoria de Guadalajara, Jalisco, México. *Salud mental*, 28(5), 64-70.
4. Inglés, C. J., Delgado, B., Bautista, R., Torregrosa, M. S., Espada, J. P., García-Fernández, J. M., ... & García-López, L. J. (2007). Factores psicosociales relacionados con el consumo de alcohol y tabaco en adolescentes españoles.
5. Nebot, M., Tomás, Z., Ariza, C., Valmayor, S., López, M. J., & Juárez, O. (2004). Factores asociados con el inicio del tabaquismo: seguimiento a los 3 años de una cohorte de escolares. *Archivos de Bronconeumología*, 40(11), 495-501.
6. OMS.(2018).Tabaco.09/03,de OMS Sitio web: <https://www.who.int/es/news-room/fact-sheets/detail/tobacco>
7. Cano, A. J., Solanas, S. E., Marí-Klose, M., & Marí-Klose, P. (2012). Factores de riesgo psicosociales en el consumo de tabaco de los adolescentes: estados de ánimo negativos, grupo de iguales y estilos parentales. *Adicciones*, 24(4), 309-317.