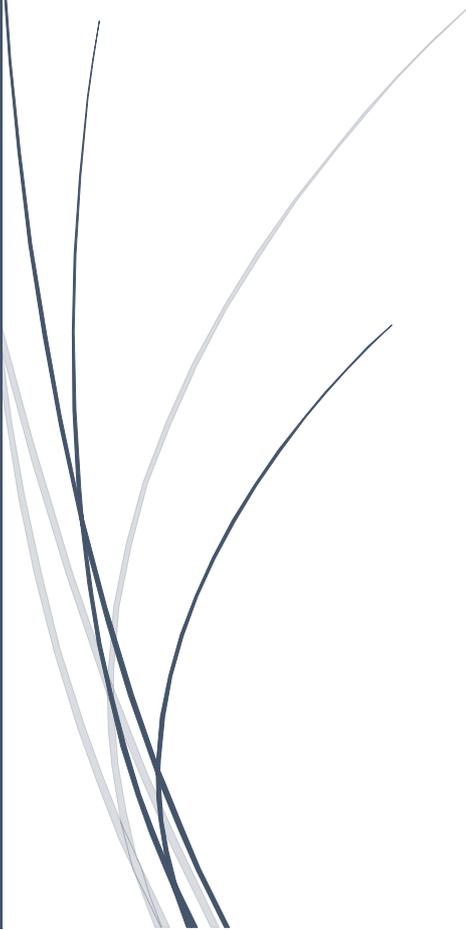


A dark blue vertical bar on the left side of the page. A blue arrow-shaped graphic points to the right from the bar, containing the date.

16-4-2021

# Ludopatía y el consumo de juegos de azar

Econometría II

Several thin, curved lines in shades of blue and grey originate from the bottom left corner and extend upwards and to the right, creating a decorative, abstract pattern.

Daniel Ribas Gómez y Lutero Vogl Ramiro  
3º GECO A

## 0 TABLA DE CONTENIDO

---

1	Motivación .....	2
2	Especificación del Modelo.....	3
3	Estimación del modelo .....	8
3.1	Detección multicolinealidad.....	9
3.1.1	Coeficientes de correlación, usando las observaciones 1 – 253 .....	9
3.1.2	VIF y Número de Condición.....	9
3.2	Soluciones a la multicolinealidad .....	10
3.2.1	Centrar variables .....	10
3.2.2	Variables ortogonales.....	12
	Estimación del modelo Logit con la variable ortogonal .....	13
4	Interpretación de los coeficientes ( $\beta_j$ ) .....	15
5	Odds, odds ratio y efectos marginales .....	15
5.1	Odds: $o_i = e^{z_i}$ .....	15
5.2	Efectos marginales .....	16
5.3	Odd ratios.....	17
6	Problemáticas.....	17
6.1	PROBLEMÁTICA MUESTRAL .....	17
6.2	Porcentaje de la población con Y=1 .....	18
7	Conclusión .....	19
	Bibliografía .....	19

### **NOTA IMPORTANTE SOBRE LOS DATOS**

- EN EL ARCHIVO ZIP EN EL QUE SE ENCONTRABA ESTE DOCUMENTO, HABÍA DOS HOJAS DE CÁLCULO EXCEL. LA PRIMERA CONTIENE LA MUESTRA COMPLETA Y CORREGIDA OBTENIDA EN LA ENCUESTA. LA SEGUNDA, LA REDUCCIÓN DE LOS DATOS UTILIZADA EN EL [apartado 7.2](#).
- NÓTESE QUE LA VARIABLE EXPLICATIVA 'INGRESOS ANUALES' SERÁ EXCLUIDA EN LA ESTIMACIÓN DEL MODELO DEBIDO A CAUSAS QUE SE EXPONDRÁN A LO LARGO DE ESTE INFORME.
- TAMBIÉN PUEDE ACCEDER A LOS RESULTADOS DE LA ENCUESTA EN EL SIGUIENTE ENLACE:  
<https://docs.google.com/forms/d/1WoC84EYkHRcV7-6gXZ21-dTd6ShIXEYPPBHp7peAK38/edit?usp=sharing>



# 1 MOTIVACIÓN

---

El lector se encuentra ante un problema que la sociedad viene arrastrando desde las primeras civilizaciones del mundo: la ludopatía. Si bien nuestro estudio no está plenamente enfocado en dar una explicación a la existencia de esta patología, se pretende ofrecer una respuesta a qué factores pueden afectar a que la población actual pueda ser más o menos propensa a consumir los juegos de azar. Más adelante se explicará el enfoque del estudio.

Los historiadores datan el origen de los juegos de azar hace más de cuatro mil años, en las poblaciones Sumeria y Asiria<sup>1</sup>, que serán las precursoras de lo que hoy conocemos como los dados. Se conoce que los emperadores romanos Augusto y Claudio llegaron a desarrollar una fuerte adicción por este tipo de juegos. Con el paso de los años, serán las clases altas las más involucradas en este tipo de prácticas dadas sus congregaciones en villas de campo -*casinò* en italiano-. Por este motivo, una vez se crean los casinos situados en las ciudades como establecimientos para el consumo específico de los juegos de azar -Venecia, s. XVII-<sup>2</sup>, podemos observar que existe una concepción de lujo, desorden y embriaguez sobre las personas que asisten a este tipo de lugares. En este sentido, podemos identificar dicha ideología en la narración sobre el casino que el mismo F. Dostoyevski presenta en su libro *El jugador* (1867)<sup>3</sup>, en el que sitúa los personajes más importantes de la aristocracia rusa, francesa e inglesa del s. XIX derrochando su fortuna en la ruleta. De esta forma, hasta bien asentado el siglo XX, la sociedad no ha relacionado el consumo de los juegos de azar con el juego patológico, la ludopatía. Es más, hoy día podemos encontrar una percepción bastante amplia de lo inofensivo de estos juegos.

A pesar del conocimiento de estas costumbres, en nuestra opinión, el factor determinante de la expansión en el consumo de los juegos de azar llega junto con la liberalización de estos servicios mediante las plataformas on-line y la informatización de los salones de juego, que permiten el acceso a cualquier persona mayor de dieciocho años. Sin embargo, el hecho de que, en España, los juegos de azar con mayor capacidad de adicción estaban prohibidos hasta 1977, y que la oferta de juego legal se limitara a la Lotería Nacional, el cupón de la ONCE y las apuestas futbolísticas (Ibáñez, A., Saiz, J., 2001), puede ser otra de las cuestiones que expliquen este incremento en el consumo. Así pues, las nuevas formas de juego afectan, sobre todo, a los más vulnerables: los jóvenes<sup>4</sup>. En esta línea, el periódico ABC publicaba en septiembre de 2019 un artículo que señala que España cuenta con la tasa más alta de Europa de ludópatas entre 14 y 21 años<sup>5</sup>. No obstante, según un estudio de la Asociación Gallega de Ludópatas Rehabilitados (Agalure) expuesto en el periódico La Vanguardia<sup>6</sup>, el perfil del ludópata en 2019 es el de hombre de 36 años, con educación secundaria y trabajo fijo.

---

<sup>1</sup> RedHistoria - *Los orígenes de los juegos de azar*

(<https://redhistoria.com/los-origenes-de-los-juegos-de-azar/#:~:text=Los%20primeros%20juegos%20de%20azar,diferentes%2C%20siendo%20el%20precursor%20del>)

<sup>2</sup> Casi-casino - *Breve historia de los casinos*

(<https://www.casi-casino.com/historia-de-los-casinos/#:~:text=El%20primer%20casino%20de%20la,muy%20parecido%20al%20Black%2DJack,>)

<sup>3</sup> Dostoyevski, F. - *El jugador* ([https://es.wikisource.org/wiki/El\\_jugador](https://es.wikisource.org/wiki/El_jugador))

<sup>4</sup> Palumberi, E., Mannino, G. - *Ludopatía. Un estudio comparativo realizado en Italia y España* (<http://papelesdesociedad.info/IMG/pdf/67811462007.pdf>)

<sup>5</sup> Periódico ABC - *España cuenta con la tasa más alta de Europa de ludópatas entre 14 y 21 años*

([https://www.abc.es/familia/padres-hijos/abci-espana-cuenta-tasa-mas-alta-europa-ludopatas-entre-14-y-21-anos-201909110151\\_noticia.html](https://www.abc.es/familia/padres-hijos/abci-espana-cuenta-tasa-mas-alta-europa-ludopatas-entre-14-y-21-anos-201909110151_noticia.html))

<sup>6</sup> La Vanguardia - *El perfil del ludópata en 2019 es el de hombre de 36 años, con educación secundaria y trabajo fijo*

(<https://gacetamedica.com/investigacion/el-perfil-del-ludopata-en-2019-es-el-de-hombre-de-36-anos-con-educacion-secundaria-y-trabajo-fijo-ae2237736/>)

## 2 ESPECIFICACIÓN DEL MODELO

---

Como ya hemos adelantado previamente, el propósito de nuestro estudio no es explicar el juego patológico -dada la profundidad que podría alcanzar este tipo de búsqueda-, sino determinar qué factores pueden determinar la propensión al consumo de los juegos de azar. Con este propósito, hemos creado una encuesta a ser respondida por nuestro círculo de contacto, conformado por nuestros familiares, amigos o conocidos. De dicha encuesta hemos extraído la información correspondiente con la variable explicada y las variables explicativas. Para mantener cierta claridad en la exposición del trabajo, mostraremos la encuesta conforme vayamos definiendo las variables con las que hemos tratado.

La muestra con la que, finalmente, hemos trabajado tiene un tamaño de 254 observaciones. Se trata de datos de corte transversal, puesto que el estudio se centra en el consumo de juegos de azar que los participantes han hecho *últimamente*, sin especificar el progreso que el consumo podría haber tenido en el tiempo.

### VARIABLE EXPLICADA

En primer lugar, nuestra variable explicada, de naturaleza binaria y de corte transversal, responde a si el individuo en cuestión ha consumido últimamente (o no) algunos de los juegos de azar más conocidos o adictivos entre la sociedad actual. Bajo este criterio, hemos decidido preguntar a los participantes sobre aquellos que mantienen las dos cualidades. De este modo, según Ibáñez, A. y Saiz, J. -entre uno de los diversos estudios que lo señalan-, los juegos más adictivos son “aquellos con un intervalo menor entre la apuesta y el premio”. De acuerdo con esta definición, hemos elegido los salones de juego -sin especificar el servicio que se consuma dentro de los mismos-, plataformas de apuestas on-line y de juego (póker u otros) -bajo el criterio de ser uno de los más conocidos actualmente- o máquinas de azar (*tragaperras*, recreativas o de arcade). Por lo tanto, consideramos que el individuo consume juegos de azar ( $y=1$ ) si responde cualquier opción excepto la que versa *No he consumido ninguno de estos servicios últimamente*, en cuyo caso consideraríamos  $y=0$ .

¿Ha consumido usted alguno de estos servicios últimamente?

- Salones de juego
- Plataformas de apuesta on-line
- Plataformas de juego (poker u otros)
- Máquinas de azar (tragaperras, recreativas, arcade...)
- Varios de estos servicios
- No he consumido ninguno de estos servicios últimamente

### VARIABLES EXPLICATIVAS

Para la selección de nuestras variables explicativas nos hemos basado, fundamentalmente, en un estudio realizado por el Ministerio de Consumo sobre prevalencia, comportamiento y características de los usuarios de juegos de azar en España realizado en 2015.

De entre las variables que ha tenido en consideración este análisis para determinar el perfil del jugador patológico, hemos seleccionado el género, la edad, el nivel de ingresos -anual en nuestro caso- y el estado civil (si tiene pareja o no). Bajo nuestro propio criterio, hemos decidido indagar en otras cuestiones: si el individuo tiene hijos, el nivel de estudios, si consume alcohol, tabaco o algún tipo de estupefaciente y si el individuo practica deporte.

- **Género**

Hemos concebido que, a priori, es una de las características que podrían resultar más relevantes a la hora de explicar la propensión a consumir juegos de azar. La naturaleza de esta variable es discreta binaria, puesto que distinguimos entre *Varón* o *Mujer*, de tal forma que si el participante responde *Varón*,  $x=1$ ; mientras que si responde *mujer*, el valor sea  $x=0$ .

Cabe destacar que hemos especificado una opción en la que el individuo pueda elegir el género con el que se identifica. El resultado han sido dos respuestas nulas que hemos decidido borrar dada su intención de burla a nuestro propósito inicial, que era mantener la tolerancia en la realización del estudio.

Basándonos en el -ya mencionado anteriormente- estudio del Ministerio de Consumo, esperamos que la propensión a consumir sea más alta para los hombres que para las mujeres, puesto que el perfil del jugador patológico se corresponde con el de un varón.

Seleccione su género \*

Varón

Mujer

Otra...

- **Edad**

Al igual que el género, la edad parece ser una de las cualidades que pueden tener una mayor implicación en la explicación de nuestra variable dependiente. En este caso, estamos ante una variable continua, ya que, aunque nuestro ámbito de estudio sea reducido, esperamos respuestas de individuos de cualquier edad, tomando valores positivos.

Según el análisis del Ministerio de consumo, cabría esperar que los individuos entre 35 y 50 años sean los más propensos a consumir.

Indique su edad \*

Texto de respuesta corta

- **Estado civil**

Podemos deducir que el estado civil de una persona puede ser relevante a la hora de determinar la propensión que un individuo pueda tener a consumir los juegos de azar en el sentido de que tener pareja aumenta las responsabilidades -sea de la forma que sea- en la vida de la persona. En esta línea, hemos tratado de definir el estado civil sin profundizar en su naturaleza, simplemente si el participante tiene pareja o no en el momento en que realiza la encuesta. Por lo tanto, la naturaleza de la variable es discreta binaria, tomando un valor  $x=1$  si el individuo no tiene pareja y  $x=0$  si el individuo tiene pareja.

Los estudios apuntan que son las personas solteras las que tienen mayor propensión al consumo de los juegos de azar.

¿Actualmente tiene pareja? \*

Sí

No

- **Nivel de estudios**

En este caso, hemos determinado que el nivel de estudios de una persona podría ser de interés gracias a que la mayor parte de nuestro ámbito de estudio recoge a personas universitarias. En segundo lugar, como se ha señalado en la motivación de nuestro estudio, en España nos encontramos ante una alarmante situación de adicción al juego entre los jóvenes, por lo que nuestro análisis nos permitiría determinar si los jóvenes con estudios son más o menos propensos al consumo que los jóvenes sin estudios. Por último, también es de valorar si podemos esperar que una persona con estudios sea más consciente de sus actos y, en consecuencia, menos propensa a consumir este tipo de servicios con alta probabilidad de adicción.

La variable en cuestión es de naturaleza discreta binaria, puesto que nuestro interés principal reside en si el participante posee estudios universitarios o de otra índole. Es de indicar que hemos considerado que las personas que actualmente estén cursando un grado universitario también pueden situarse en el grupo de personas con estudios universitarios. En este sentido, la variable tomará un valor de  $x=1$  si el individuo no tiene estudios universitarios y  $x=0$  en caso contrario.

¿Cuál es su nivel de estudios? \*

Si actualmente está cursando un grado universitario, marque la opción "Universitarios".

Universitarios

No universitarios

- **Nivel de ingresos anual**

Hemos considerado los ingresos anuales (en euros) como una de las variables que mejor puede explicar que un individuo sea más o menos propenso a consumir juegos de azar. Esta variable guarda una naturaleza continua, ya que esperamos valores positivos y sin restricción

superior. Particularmente, hemos propuesto como válida cualquier fuente de ingresos, teniendo en cuenta que gran parte de las respuestas podrían ser de estudiantes que dependen de sus progenitores para mantenerse, indicando que las becas o las transferencias familiares también son aceptables.

En base al estudio del Ministerio de Consumo, el perfil del jugador patológico mantiene unos ingresos mensuales de entre 600 y 1500 euros mensuales, lo que se traduce en unos ingresos anuales de entre 7.200 y 18.000 euros en nuestro modelo.

En el caso de esta variable hemos decidido, al igual que con la variable género, descartar los resultados que parecen no tener coherencia con lo que esperamos de nuestro ámbito de estudio. Por ejemplo, un millón quinientos mil euros como respuesta no se corresponde con nuestra realidad esperada.

Indique su nivel de ingresos ANUAL de forma aproximada \*

Es válida cualquier tipo de fuente de ingresos: becas, transferencias familiares, remuneración del trabajo, etc.

Texto de respuesta corta

- **Consumo de alcohol**

El consumo de alcohol ha sido una de las variables que hemos decidido bajo criterio propio, con tal de averiguar si podemos esperar alguna relación entre el consumo habitual de alcohol y el consumo de juegos de azar.

Con esta meta hemos redactado una pregunta que analice la frecuencia con la que los participantes consumen alcohol, proponiendo como posibles respuestas *consumo eventual -en citas o celebraciones-*, *consumo en fines de semana*, *más de tres días por semana* o *no consumo*. Para facilitar el análisis hemos transformado este abanico de respuestas en nuestro interés, que es si los individuos, efectivamente, consumen alcohol o no. De esta forma, consideramos como consumo de alcohol las personas que consumen los fines de semana o más de tres días por semana; mientras que consideramos más razonable y menos abusivo el hecho de que un individuo consuma alcohol en celebraciones o citas puntuales o, simplemente, que no consuma alcohol. Por lo tanto,  $x=1$  cuando consume frecuentemente, es decir, cuando los individuos responden *Fines de semana* o *Más de tres días por semana*; y  $x=0$  cuando no se consume frecuentemente, o sea, cuando el participante responde *Eventualmente (celebraciones o citas puntuales)* o *No consumo alcohol*.

Podríamos presumir que, aquellas personas que mantienen un consumo regular de alcohol pueden ser más propensas a consumir juegos de azar, ya que la bebida altera el estado natural de los individuos.

Indique con qué frecuencia consume alcohol \*

- Eventualmente (celebraciones o citas puntuales)
- Fines de semana
- Más de tres días por semana
- No consumo alcohol

- **Consumo de tabaco o algún tipo de estupefaciente**

El consumo de tabaco o estupefacientes nos resulta de interés, al igual que el alcohol, gracias a que pueden llegar a alterar la conducta de los individuos cuando se trata de un consumo abusivo o habitual. En este caso, hemos propuesto unas respuestas simples que determinan la naturaleza discreta binaria de la variable en cuestión: *¿Consume usted de manera frecuente tabaco o algún tipo de estupefaciente?*. Así pues, diremos que  $x=1$  si el individuo consume y  $x=0$  en caso contrario.

Podríamos deducir que, si el individuo consume tabaco o algún tipo de estupefaciente podría ser más vulnerable a otro tipo de adicciones como puede ser, en este caso, el consumo de juegos de azar.

¿Consume usted de manera frecuente tabaco o algún tipo de estupefaciente? \*

Sí

No

- **Deporte**

De la misma forma que con el consumo de alcohol, hemos traducido la frecuencia con la que los individuos practican deporte en si, efectivamente, lo practican o no. De esta forma, hablamos de una variable discreta binaria. En este caso, hemos considerado que las personas practican deporte de forma habitual si lo hacen más de un día por semana. Por lo tanto, si el individuo responde *No practico deporte* o *Un día a la semana*, consideraremos  $x=1$ ; mientras que si el participante responde cualquiera de las otras opciones disponibles, la variable tomará un valor de  $x=0$ .

Sería de esperar que las personas que invierten parte del día en mantenerse sanos físicamente -personas que practican deporte-, sean menos vulnerables a caer en el consumo de juegos de azar.

Indique con qué frecuencia practica deporte \*

No practico deporte

Un día a la semana

Entre uno y tres días a la semana

Entre tres y cuatro días por semana

Más de cuatro días por semana

- **Hijos**

De nuevo, hemos propuesto diversas respuestas que, posteriormente, han sido traducidas en una variable binaria, ya que nuestro objetivo era determinar si la responsabilidad que supone tener hijos -independientemente de cuántos sean- influye a la hora de ser más o menos propenso a consumir juegos de azar. Así pues, cuando los individuos responden *No tengo hijos*,  $x=0$  y  $x=1$  en el resto de respuestas posibles.

En este caso, deducimos que las personas que no tienen hijos tienen menos responsabilidades y, por tanto, pueden tener una propensión mayor a consumir juegos de azar.

Indique si posee algún hijo \*

- No tengo hijos
- Tengo un hijo
- Tengo dos hijos
- Tengo tres o más hijos



### 3 ESTIMACIÓN DEL MODELO

Modelo 1: Logit, usando las observaciones 1-253

Variable dependiente: y

Desviaciones típicas basadas en el Hessiano

	coeficiente	Desv. típica	z	valor p	
const	-1.77822	0.798762	-2.226	0.0260	**
Genero	0.891432	0.437225	2.039	0.0415	**
Edad	-0.0379716	0.0293786	-1.292	0.1962	
Pareja	0.0263395	0.413211	0.06374	0.9492	
Niv_estudios	0.624138	0.463676	1.346	0.1783	
Cons_Alcohol	0.486884	0.415330	1.172	0.2411	
Tabaco_estupef	0.262793	0.515102	0.5102	0.6099	
Frec_deporte	0.0658992	0.438506	0.1503	0.8805	
Hijos_b	-0.186606	1.07014	-0.1744	0.8616	
Media de la vble. dep.	0.130435	D.T. de la vble. dep.	0.337449		
R-cuadrado de McFadden	0.090950	R-cuadrado corregido	-0.000920		
Log-verosimilitud	-89.05489	Criterio de Akaike	196.1098		
Criterio de Schwarz	227.9103	Crit. de Hannan-Quinn	208.9042		

Número de casos 'correctamente predichos' = 220 (87.0%)

f(beta\*x) en la media de las variables independientes = 0.092

Contraste de razón de verosimilitudes: Chi-cuadrado(8) = 17.8197 [0.0226]

	Predicho		
	0	1	
Observado 0	220	0	
1	33	0	

Sin considerar la constante, el valor p más alto fue el de la variable 4 (Pareja)

En primer lugar, hemos estimado el modelo mediante un modelo Logit incluyendo todas las variables independientes que hemos considerado relevantes. No obstante, podemos observar que la única variable explicativa que en principio parece significativa es 'Género'.

Nótese que no aparece la variable 'Ingresos Anuales'. Hemos tomado esta decisión, en primer lugar, debido a los errores cometidos a la hora de formular la encuesta (véase el apartado [6](#)).

[problemáticas](#)) y, en segundo lugar, debido a la ausencia de significación individual no solo al estimarla en el modelo anterior junto a las demás variables sino de forma independiente. Es decir, en la estimación del siguiente modelo,  $\alpha_1$  no es significativa.

$$y_i = \frac{e^{z_i}}{1+e^{z_i}} + u_i; \text{ donde } z_i = \alpha_0 + \alpha_1 * (\text{Ingresos Anuales})$$

### 3.1 DETECCIÓN MULTICOLINEALIDAD

#### 3.1.1 Coeficientes de correlación, usando las observaciones 1 – 253

Valor crítico al 5% (a dos colas) = 0.1234 para n = 253

Genero	Edad	Pareja	Niv_estudios	Cons_Alcohol	
1.0000	-0.1234	0.1327	0.0346	0.0772	<b>Genero</b>
	1.0000	-0.2532	0.3127	0.0101	<b>Edad</b>
		1.0000	-0.0976	0.0359	<b>Pareja</b>
			1.0000	-0.0905	<b>Niv_estudios</b>
				1.0000	<b>Cons_Alcohol</b>
		<b>Tabaco_estup</b>	<b>Frec_deporte</b>	<b>Hijos_b</b>	
		0.1030	-0.2329	-0.1077	<b>Genero</b>
		-0.1089	0.0353	<b>0.7582</b>	<b>Edad</b>
		0.0590	-0.0430	-0.2657	<b>Pareja</b>
		0.0165	-0.0147	0.3671	<b>Niv_estudios</b>
		0.2532	-0.0075	0.0198	<b>Cons_Alcohol</b>
		1.0000	0.0985	-0.1224	<b>Tabaco_estup</b>
			1.0000	0.0789	<b>Frec_deporte</b>
				1.0000	<b>Hijos_b</b>

Aquí podemos observar que existe una correlación bastante elevada entre 'Edad' e 'Hijos\_b'.

#### 3.1.2 VIF y Número de Condición

Factores de inflación de varianza (VIF)

Mínimo valor posible = 1.0

Valores mayores que 10.0 pueden indicar un problema de colinealidad

Genero	1.107
Edad	2.392
Pareja	1.096
Niv_estudios	1.195
Cons_Alcohol	1.096
Tabaco_estupef	1.125
Frec_deporte	1.085
Hijos_b	2.541

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

Diagnósticos de colinealidad de Belsley-Kuh-Welsch:

proporciones de la varianza

lambda	cond	const	Genero	Edad	Pareja	Niv_estu~	Cons_Alc~	Tabaco_e~
4.836	1.000	0.002	0.009	0.002	0.010	0.011	0.012	0.009
1.167	2.036	0.000	0.004	0.003	0.032	0.051	0.014	0.071
0.834	2.408	0.001	0.041	0.000	0.074	0.072	0.017	0.181
0.666	2.695	0.002	0.001	0.002	0.038	0.401	0.015	0.362
0.615	2.804	0.000	0.002	0.001	0.021	0.096	0.317	0.025
0.376	3.588	0.001	0.157	0.003	0.001	0.365	0.505	0.192
0.311	3.942	0.006	0.207	0.002	0.720	0.004	0.014	0.134
0.164	5.428	0.095	0.554	0.085	0.048	0.001	0.106	0.024
0.031	12.480	0.893	0.025	0.902	0.055	0.000	0.001	0.001

lambda	cond	Frec_dep~	Hijos_b
4.836	1.000	0.010	0.003
1.167	2.036	0.001	0.155
0.834	2.408	0.229	0.017
0.666	2.695	0.062	0.003
0.615	2.804	0.386	0.023
0.376	3.588	0.042	0.036
0.311	3.942	0.042	0.112
0.164	5.428	0.200	0.153
0.031	12.480	0.028	0.498

lambda = autovalores de la inversa de la matriz de covarianzas (smallest is 0.0310509)

cond = índice de condición

nota: Las columnas de proporciones de la varianza suman 1.0

De acuerdo con BKW, cond  $\geq 30$  indica "fuerte" dependencia casi lineal, y cond entre 10 y 30 "moderadamente fuerte".

---

No parece haber ningún FIV que indique colinealidad. Pero el número de condición (NC) si la indica, ya que podemos observar un valor superior a 10 (12.480) que es indicativo de la existencia de una colinealidad 'moderadamente fuerte'.

Si bien estos dos indicadores parecen contradecirse, esto es totalmente factible debido a que, a diferencia del NC, el FIV es únicamente capaz de detectar la multicolinealidad esencial. Por ello, es probable que exista algún problema de multicolinealidad no esencial. Es decir, debido a la relación de la constante con el resto de variables independientes.

## 3.2 SOLUCIONES A LA MULTICOLINEALIDAD

### 3.2.1 Centrar variables

Una posible solución a la multicolinealidad no esencial, es centrar las variables que tengan una varianza relativamente pequeña. En este caso, para eliminarla hemos centrado a la variable continua 'Edad'

$$Edad\_centrada = Edad - \text{mean}(Edad) = Edad - 30.7945$$

Modelo 2: Logit (observaciones 1-253), con variable 'Edad' centrada

Variable dependiente: y  
Desviaciones típicas basadas en el Hessiano

	coeficiente	Desv. típica	z	valor p
const	-2.94754	0.549603	-5.363	8.18e-08 ***
Genero	0.891432	0.437225	2.039	0.0415 **
Pareja	0.0263395	0.413211	0.06374	0.9492
Niv_estudios	0.624138	0.463476	1.346	0.1783
Cons_Alcohol	0.486884	0.415330	1.172	0.2411
Tabaco_estupef	0.242793	0.515102	0.5102	0.6099
Frec_deporte	0.0655992	0.438506	0.1503	0.8805
Hijos_b	-0.186606	1.07014	-0.1744	0.8616
Edad_centrada	-0.0379716	0.0293786	-1.282	0.1962

Media de la vble. dep. 0.130435 D.T. de la vble. dep. 0.337449  
R-cuadrado de McFadden 0.090950 R-cuadrado corregido -0.000920  
Log-verosimilitud -89.05489 Criterio de Akaike 196.1098  
Criterio de Schwarz 227.9103 Crit. de Hannan-Quinn 208.9042

Número de casos 'correctamente predichos' = 220 (87.0%)  
f(beta'x) en la media de las variables independientes = 0.092  
Contraste de razón de verosimilitudes: Chi-cuadrado(8) = 17.8197 [0.0226]



Observado	Predicho	
	0	1
0	220	0
1	33	0

Sin considerar la constante, el valor p más alto fue el de la variable 4 (Pareja)

(...)

Diagnósticos de colinealidad de Belsley-Kuh-Welsch:

proporciones de la varianza

lambda	cond	const	Genero	Pareja	Niv_estu~	Cons_Alc~	Tabaco_e~	Frec_dep~
4.235	1.000	0.006	0.012	0.015	0.012	0.015	0.014	0.013
1.634	1.610	0.000	0.000	0.007	0.035	0.001	0.001	0.002
0.854	2.227	0.001	0.028	0.040	0.134	0.038	0.246	0.173
0.654	2.545	0.004	0.007	0.022	0.057	0.053	0.332	0.392
0.631	2.590	0.002	0.005	0.022	0.366	0.281	0.094	0.089
0.371	3.377	0.001	0.150	0.039	0.372	0.485	0.202	0.027
0.326	3.607	0.018	0.204	0.812	0.009	0.001	0.067	0.003
0.211	4.485	0.039	0.346	0.011	0.006	0.018	0.037	0.208
0.085	7.075	0.928	0.247	0.033	0.009	0.108	0.007	0.092

lambda	cond	Hijos_b	Edad_cen~
4.235	1.000	0.001	0.004
1.634	1.610	0.096	0.061
0.854	2.227	0.000	0.003
0.654	2.545	0.008	0.005
0.631	2.590	0.012	0.007
0.371	3.377	0.047	0.007
0.326	3.607	0.001	0.049
0.211	4.485	0.393	0.416
0.085	7.075	0.442	0.448

(...)

No evidence of excessive collinearity

Tras calcular el nuevo modelo, vemos que se ha conseguido reducir el NC a 7.075(<10). Eliminando así el problema de multicolinealidad no esencial. Además, podemos observar cómo la estimación de los parámetros de las variables independientes (a excepción de  $\beta_0$ ) no varía, independientemente de si se utilice la variable 'Edad' o su centrada.

No obstante, centrar las variables no es un remedio útil para la multicolinealidad esencial (probado que no existe en nuestro caso con el FIV) o sistémica. Así, el coeficiente de correlación simple entre las variables 'Edad\_centrada' e 'Hijos\_b' sigue siendo exactamente el mismo que cuando la primera aún no se había centrado (0.7582).

### 3.2.2 Variables ortogonales

Una posible solución al problema de correlación entre 'Edad\_centrada' e 'Hijos\_b' puede ser mediante la utilización de variables ortogonales. Para ello comenzamos estimando mediante **EMCO** la regresión auxiliar de 'Edad\_centrada' (variable continua) con respecto a la constante, 'Hijos\_b' y 'Género' ya que las demás variables explicativas no son significativas en este modelo.

#### Regresión auxiliar

Modelo 3: MCO (observaciones 1-253), Regresión auxiliar

Variable dependiente: Edad\_centrada

	coeficiente	Desv. típica	Estadístico t	valor p	
const	-6.30570	1.05160	-5.996	7.03e-09	***
Genero	-1.42090	1.39152	-1.021	0.3082	
Hijos_b	29.9054	1.64243	18.21	9.87e-048	***
Media de la vble. dep.	-0.000034	D.T. de la vble. dep.		16.81285	
Suma de cuad. residuos	30156.61	D.T. de la regresión		10.98301	
R-cuadrado	0.576650	R-cuadrado corregido		0.573263	
F(2, 250)	170.2641	Valor p (de F)		2.17e-47	
Log-verosimilitud	-963.7588	Criterio de Akaike		1933.518	
Criterio de Schwarz	1944.118	Crit. de Hannan-Quinn		1937.783	

Una vez obtenidos los parámetros estimados, buscamos los residuos creando una nueva variable. Estos residuos obtenidos en la regresión auxiliar serán introducidos en la regresión ortogonal, dejando fuera de la regresión a 'Edad\_centrada'. Así, el modelo a estimar quedaría de la siguiente forma:

$$y_i = \frac{e^{z_i}}{1+e^{z_i}} + u_i$$

Donde  $z_i = \beta_0 + \beta_1 * \text{Género} + \beta_2 * (\text{Hijos}_b) + \beta_3 * (\text{residuos\_Edad\_estimados})$

## Estimación del modelo Logit con la variable ortogonal

Modelo 4: Logit (observaciones 1-253), estimación con variable ortogonal

```

Variable dependiente: y
Desviaciones típicas basadas en el Hessiano

-----
                coeficiente  Desv. típica    z      valor p
-----
const          -2.35874      0.352820    -6.688  2.30e-011 ***
Genero         1.05335      0.410460     2.566  0.0103 **
Hijos_b       -1.11707      0.632801    -1.765  0.0775 *
residuos_Edad_-_2  -0.0344248    0.0280881   -1.226  0.2203

Media de la vble. dep.  0.130435  D.T. de la vble. dep.  0.337449
R-cuadrado de McFadden  0.071191  R-cuadrado corregido  0.030360
Log-verosimilitud     -90.99056  Criterio de Akaike    189.9811
Criterio de Schwarz   204.1147  Crit. de Hannan-Quinn 195.6675

Número de casos 'correctamente predichos' = 220 (87.04)
f(beta*x) en la media de las variables independientes = 0.085
Contraste de razón de verosimilitudes: Chi-cuadrado(3) = 13.9483 [0.0030]

      Predicho
      0      1
Observado 0  220  0
          1   33  0

Sin considerar la constante, el valor p más alto fue el de la variable 17 (residuos_Edad_centrada_2)

```

Coeficientes de correlación, usando las observaciones 1 - 253

Valor crítico al 5% (a dos colas) = 0.1234 para n = 253

<b>Hijos_b</b>	<b>residuos_Edad_c entrada_2</b>	
1.0000	0.0000	<b>Hijos_b</b>
	1.0000	<b>residuos_Edad_c entrada_2</b>

Podemos observar como aplicando el modelo de regresión de **Wooldridge (2006)** con variables ortogonales, se ha conseguido eliminar la correlación entre Hijos y Edad. Por otro lado, es importante observar que independientemente de que utilicemos la variable 'Edad\_centrada' o su ortogonal 'Residuos\_edad\_centrada' su coeficiente estimado es idéntico.

No obstante, una vez eliminados todos los indicios de multicolinealidad, seguimos encontrándonos con un problema. Por un lado, como hemos visto en el cuadro anterior, al estimar el modelo Logit con las tres variables explicativas a la vez (Género, hijos\_b y residuo de la edad), sale como variable significativa la posesión de hijos y no el residuo de la edad (*Vease modelo 4, estimación con variable ortogonal*)

Sin embargo, al analizar estas dos por separado junto con el Género, como vemos en los siguientes dos cuadros de resultados, ambas aparecen como variables significativas. Por último, el segundo modelo, que muestra la variable explicativa 'Edad' como significativa, presenta un contraste de razón de verosimilitud y un R<sup>2</sup> de McFadden superior que el primero, por ello escogemos este como el modelo más significativo y descartamos la variable explicativa 'Hijos\_b'.



Además, observamos que, en ambos modelos, el **contraste de la razón de verosimilitud** rechaza la hipótesis nula. Por ende, los coeficientes de ambos modelos son significativos de forma conjunta.

---

Modelo 5: Logit, usando las observaciones 1-253 (Excluyendo la variable 'Edad')

Variable dependiente: y

Desviaciones típicas basadas en el Hessiano

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>valor p</i>	
const	-2.30086	0.339477	-6.778	<0.0001	***
Genero	1.02484	0.405505	2.527	0.0115	**
<b>Hijos_b</b>	<b>-1.12267</b>	<b>0.630423</b>	<b>-1.781</b>	<b>0.0749</b>	<b>*</b>
Media de la vble. dep.	0.130435		D.T. de la vble. dep.	0.337449	
<b>R-cuadrado de McFadden</b>	<b>0.061236</b>		R-cuadrado corregido	0.030613	
Log-verosimilitud	-91.96573		Criterio de Akaike	189.9315	
Criterio de Schwarz	200.5316		Crit. de Hannan-Quinn	194.1963	

Número de casos 'correctamente predichos' = 220 (87.0%)

f(beta'x) en la media de las variables independientes = 0.337

Contraste de razón de verosimilitudes: Chi-cuadrado (2) = 11.998 [0.0025]

---

Modelo 6: Logit, usando las observaciones 1-253 (Excluyendo la variable Hijos b)

Variable dependiente: y

Desviaciones típicas basadas en el Hessiano

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>valor p</i>	
const	-1.47952	0.547468	-2.702	0.0069	***
Genero	1.00680	0.406607	2.476	0.0133	**
<b>Edad</b>	<b>-0.0363278</b>	<b>0.0172297</b>	<b>-2.108</b>	<b>0.0350</b>	<b>**</b>
Media de la vble. dep.	0.130435		D.T. de la vble. dep.	0.337449	
<b>R-cuadrado de McFadden</b>	<b>0.071154</b>		R-cuadrado corregido	0.040531	
Log-verosimilitud	-90.99414		Criterio de Akaike	187.9883	
Criterio de Schwarz	198.5885		Crit. de Hannan-Quinn	192.2531	

Número de casos 'correctamente predichos' = 220 (87.0%)

f(beta'x) en la media de las variables independientes = 0.337

Contraste de razón de verosimilitudes: **Chi-cuadrado (2) = 13.9412 [0.0009]**

---

Otro gran problema en el modelo es el bajo valor que tiene el pseudo R<sup>2</sup> de McFadden y el hecho de que a pesar de que el número de casos correctamente predichos es elevado (87%), esto es debido a que el modelo siempre predice que 'Y' va a ser igual a cero.

Tabla de aciertos		Predicho	
		0	1
Observado	0	220	0
	1	33	0

Probablemente esto se deba a que más del 85% de la muestra no consumiera este tipo de servicios. Este desproporcionado resultado seguramente sea causado por el tipo de observaciones que contiene la muestra. Mayoritariamente universitarios, amigos y familiares de nuestro entorno que pueden (aunque no con toda certeza) haber sesgado de forma determinante los resultados que se hubieran obtenido, si la encuesta se hubiera hecho a un grupo más diverso y representativo de la población real en nuestro país.

## 4 INTERPRETACIÓN DE LOS COEFICIENTES ( $\beta_j$ )

En los modelos Logit, a diferencia de lo que ocurre en los modelos estimados por EMCO, los coeficientes no son los efectos marginales. No obstante, los signos de estos si nos indican si su correspondiente variable explicativa tiene una relación inversa o directa con la variable explicada. Así, podemos observar una relación inversa entre la edad y la probabilidad de consumir juegos de azar ( $p_i$ ) y un aumento de  $p_i$  si el individuo es hombre.

## 5 ODDS, ODDS RATIO Y EFECTOS MARGINALES

Debido a que la muestra contiene 254 individuos y cada uno de ellos tiene su correspondiente efecto marginal y odd, únicamente mostraremos los valores de estos indicadores para el individuo típico y el máximo y mínimo de cada caso.

En todos ellos, partimos del modelo estimado mediante Logit:

$$\hat{y} = \frac{e^{z_i}}{1 + e^{z_i}}; \text{ donde } z_i = -1.47952 + 1.00680 * \text{Género} - 0.0363278 * \text{Edad}$$

### 5.1 ODDS: $o_i = e^{z_i}$

#### Máximo odd: $o_i = 0.361448$

Se encuentra en la observación 131 y es un individuo varón de 15 años que no ha consumido ningún juego de azar recientemente. Así, la interpretación de este coeficiente es que este individuo tiene 0.361448 posibilidades más de jugar a juegos de azar a no. O lo que es lo mismo, por medio de la inversa, obtenemos que el individuo tiene  $1/0.361448 = 2.77$  veces **más de probabilidades de no jugar** ( $y=0$ ) que de jugar ( $y=1$ ).



Aunque parezca contradictorio que un individuo de 15 años sea el que más **probabilidades** tenga de consumir este tipo de servicios, esto se debe al signo negativo del coeficiente de la variable independiente 'Edad'. Por ello, y debido a que solo hemos incluido dos variables explicativas, es razonable pensar que el varón más joven de la muestra (según el modelo) sea el que más **probabilidades** tenga de haber consumido estos juegos recientemente. Algo que, sin lugar a dudas, es un resultado poco creíble ya que, entre otros aspectos, **legalmente está prohibido para** menores de 18 años.

**Mínimo odd:  $\pi_i = 0.010769$**

Se encuentra en la observación 7, la cual es una mujer de 84 años que no ha consumido recientemente ninguno de estos servicios. Calculando la inversa de su  $\pi_i$ , observamos que tiene  $1/0.010769=92.86$  posibilidades **más de no consumir** estos servicios ( $Y=0$ ) que de consumirlos.

**Odd del individuo medio**

El individuo promedio será una mujer (moda de la variable 'Género') de 30.7945 años (media aritmética de la variable 'Edad') **que no ha consumido juegos de azar recientemente (moda de la variable explicada).**

Así, si calculamos su odd:

$$\pi_i = e^{-1.47952 + 1.00680 \cdot (0) - 0.0363278 \cdot (30.7954)} = 0.0744$$

Una vez más, observamos cómo es más probable que el individuo medio no consuma juegos de azar. Calculando su inversa, vemos que es **13.44 veces más probable** de que **no consuma ( $y=0$ )** a que consuma ( $y=1$ ).

## 5.2 EFECTOS MARGINALES

Variables continuas :  $em(X_{ij}) = \pi_i(1 - \pi_i) * \beta_j$

Variables discretas:  $em(X_{ij}) = P[Y = 1|X = (1, x, \dots, 1, \dots, x)] - P[Y = 1|X = (1, x, \dots, 0, \dots, x)]$

Tanto el individuo con máximo  $\pi_i$ , el de mínimo  $\pi_i$  o el individuo medio coinciden con sus respectivos del apartado anterior.

**Máximo  $\pi_i$ :  $i=131$  con el valor  $\pi_i=0.2654$**

- **Em (Edad)** =  $0.2654 \cdot (1 - 0.2654) \cdot (-0.0363278) = -0.007$
- **Em (Género)**

$$\# - z_i(\text{Género}=1) = -\{-1.47952 + 1.00680 \cdot (1) - 0.0363278 \cdot (15)\} = 1.15$$

$$\# - z_i(\text{Género}=0) = -\{-1.47952 + 1.00680 \cdot (0) - 0.0363278 \cdot (15)\} = 2.024$$

$$\text{Em (Género)} = \frac{1}{1 + e^{1.15}} - \frac{1}{1 + e^{2.024}} = \mathbf{0.12378}$$

**Mínimo  $\pi_i$ :  $i=7$  con el valor  $\pi_i= 0.01065456$**

- **Em (Edad)** =  $0.01065456 \cdot (1 - 0.01065456) \cdot (-0.0363278) = -\mathbf{0.000667}$
- **Em (Género)**

$$\# - z_i(\text{Género}=1) = -\{-1.47952 + 1.00680 \cdot (1) - 0.0363278 \cdot (84)\} = 3.43$$

$$\# - z_i(\text{Género}=0) = -\{-1.47952 + 1.00680 \cdot (0) - 0.0363278 \cdot (84)\} = 4.53$$

$$\text{Em(Género)} = \frac{1}{1+e^{3.43}} - \frac{1}{1+e^{4.53}} = 0.020705$$

### Efectos marginales del individuo medio

$$p_i = \frac{1}{1+e^{-z_i}} = 0.06925; \quad z_i = -1.47952 + 1.00680 * (\mathbf{0}) - 0.0363278 * (30.7945) = -2.5982$$

- **Em (Edad)** =  $0.06925 * (1 - 0.06925) * (-0.0363278) = -0.0023$
- **Em (Género)**
  - # - $z_i$  (Género=1) =  $-\{-1.47952 + 1.00680 * (\mathbf{1}) - 0.0363278 * (30.7945)\} = 1.166$
  - # - $z_i$  (Género=0) =  $-\{-1.47952 + 1.00680 * (\mathbf{0}) - 0.0363278 * (30.7945)\} = 2.26668$

$$\text{Em(Género)} = \frac{1}{1+e^{1.166}} - \frac{1}{1+e^{2.26668}} = 0.143$$

## 5.3 ODD RATIOS

El odd ratio representa el cambio que se produce en el odd cuando una variable explicativa,  $x_j$ , incrementa en una unidad, suponiendo que el resto de variables permanece constante. A diferencia de los *efectos marginales* y los *odds*, estos cocientes no dependen de la observación  $i$ . Por ello, únicamente existirá un odd ratio para cada coeficiente.

$$or_j = e^{\beta_j}$$

- **or<sub>Género</sub>** =  $e^{1.00680} = 2.7368$ . Es decir, suponiendo que todas las demás variables permanecen constantes, el *odd* de un hombre de consumir juegos de azar es 2.7368 veces mayor que el de una mujer. O lo que es lo mismo, el hombre tiene 2.7368 veces más posibilidades de consumir estos juegos que una mujer.
- **or<sub>Edad</sub>** =  $e^{-0.0363278} = 0.96432$ . Si calculamos su inversa (1.037), el *odd-ratio* se puede interpretar como que, por cada año adicional, las posibilidades de haber jugado a algún juego de azar recientemente son 1.037 veces menor.



## 6 PROBLEMÁTICAS

### 6.1 PROBLEMÁTICA MUESTRAL



Una vez realizada la compilación de datos y el modelo econométrico, hemos encontrado que nuestro ámbito de estudio ha sido un gran condicionante en la obtención de los datos y la realización del modelo.

En primer lugar, como en gran parte de los estudios estadísticos, debemos tener en cuenta la posibilidad de que alguno de los participantes de la encuesta haya podido resolverla de forma aleatoria, omitiendo la seriedad con la que pretendíamos dotarlo. En este sentido, hemos intentado ser lo más honestos y fieles con nuestra realidad de estudio a la hora de eliminar ciertos datos que podrían sesgar nuestro análisis. Como ya se ha indicado en la sección

de *Especificación del modelo*, hemos resuelto deshacernos de algunos datos que, claramente, mantenían una intención de burla o fallo.

Por otro lado, hemos tratado de incidir bastante en el ámbito de aplicación de nuestro estudio, pues entendemos que no podemos extrapolar nuestro análisis a la población española en general. Un ejemplo específico de esta cuestión se observa en que la mayor parte de las respuestas obtenidas en la encuesta han sido realizadas por individuos de veinte años. Esto condiciona otras preguntas de nuestra encuesta, como podrían ser el número de hijos, en la que esperamos que, por este motivo, la mayoría de participantes no tengan hijos; o el nivel de ingresos, en el que cabría presumir que los ingresos sean de un nivel de subsistencia, dado que el grupo principal de respuestas son de universitarios dependientes de sus padres, sin renta propia. Asimismo, esta ha sido la causa principal por la que hemos descartado añadir la variable trabajo en nuestro modelo.

En último lugar, como se ha mencionado anteriormente, a priori asumimos que el nivel de ingresos anual sería una de las variables más relevantes en el modelo. Esto nos lleva a una pregunta: ¿por qué ingresos anuales y no mensuales? Concluimos de nuevo que, dado que la mayor parte de nuestro ámbito de estudio se resume en estudiantes que no cuentan con ingresos propios, es más fácil indicar una cuantía aproximada anual que no una mensual, teniendo en cuenta la validez de las becas o transferencias familiares como respuesta a la encuesta. No obstante, el resultado ha sido una gran confusión entre los participantes a la hora de responder, que, en algunos casos -deducimos- han indicado sus ingresos mensuales y en otros casos han resuelto que no mantienen ninguna fuente de ingresos. Por lo tanto, esta puede ser una de las causas por las que hemos obtenido ausencia de significación individual para esta variable en el modelo conjunto -con todas las variables explicativas consideradas desde un primer momento- y, además, ausencia de significación individual en un modelo en el que solo incluimos la variable ingresos como variable explicativa. Así pues, la decisión que nos ha parecido más razonable para evitar otros sesgos en el modelo ha sido no tenerla en cuenta.

## 6.2 PORCENTAJE DE LA POBLACIÓN CON $Y=1$

Como hemos dicho anteriormente, una de las posibles causas de que nuestro modelo siempre prediga que los individuos no van a consumir juegos de azar (Véase [Tabla de aciertos](#)) es que el 86,5% de la muestra realmente no los ha consumido ( $Y=0$ ). Esto se puede deber en parte a que la encuesta ha sido mayoritariamente rellenada por familiares y amigos de nuestro entorno. Por ello, es muy probable que no sea una muestra representativa de la población.

No obstante, a pesar de no conocer la verdadera proporción de la población española que, si consume este tipo de servicios, vamos a reducir de forma arbitraria el número de individuos que no consumen juegos de azar para ver si de esta forma obtenemos mejores predicciones en nuestro modelo.

Debido a que la muestra ha sido alterada de forma no aleatoria, estos resultados no son en realidad correctos, pero nos parece interesante demostrar una de las posibles causas de la problemática anteriormente mencionada.

La siguiente muestra tiene 132 observaciones, de las cuales el 25% han consumido recientemente juegos de azar y el 75% no. Y, cómo podemos ver, el modelo ahora si predice que determinadas observaciones van a tomar el valor  $Y=1$ . Además, la  $R^2$  de Mcfadden ahora tiene un valor considerablemente superior y el modelo sigue siendo significativo en su conjunto.

Variable dependiente: y  
Desviaciones típicas basadas en el Hessiano

	coeficiente	Desv. típica	z	valor p	
const	2.08434	0.860102	2.423	0.0154	**
GÁnero	-2.62150	0.708465	-3.700	0.0002	***
Edad	-0.0336914	0.0202222	-1.666	0.0957	*
Media de la vble. dep.	0.250000	D.T. de la vble. dep.	0.434662		
R-cuadrado de McFadden	0.141727	R-cuadrado corregido	0.101311		
Log-verosimilitud	-63.70812	Criterio de Akaike	133.4162		
Criterio de Schwarz	142.0646	Crit. de Hannan-Quinn	136.9305		

Número de casos 'correctamente predichos' = 106 (80.3%)  
f(beta'x) en la media de las variables independientes = 0.175  
Contraste de razón de verosimilitudes: Chi-cuadrado(2) = 21.0402 [0.0000]

	Predicho	
	0	1
Observado 0	96	3
1	23	10

## 7 CONCLUSIÓN

El modelo que mejor explica si un individuo ha consumido recientemente algún tipo de juego de azar con la información obtenida en la encuesta es:

$$\hat{y} = \frac{e^{z_i}}{1 + e^{z_i}}; \text{ donde } z_i = -1.47952 + 1.00680 * \text{Género} - 0.0363278 * \text{Edad}$$

Podemos observar que los hombres tendrán más probabilidades de consumir este tipo de servicios que las mujeres. Además, parece ser que los jóvenes también son más propensos, lo que parece contradecir la hipótesis inicial obtenida de fuentes oficiales de que las personas más proclives a ser ludópatas tienen entre 35 y 55 años. Esto podría darse por dos razones: La primera, es que la muestra no representa realmente a la **población española**. Y la segunda, es que la tasa de ludopatía entre los jóvenes de entre 14 y 21 años es la más alta de Europa.

Por último, cabe destacar que a pesar de que el modelo es significativo en su conjunto, la R<sup>2</sup> de McFadden es muy baja y predice que ningún individuo consumirá juegos de azar. Como se ha dicho en el [apartado 6.2](#), la única forma de solucionar este problema con los datos disponibles es manipular de forma arbitraria la muestra reduciendo el número de observaciones que no consumen este servicio.

## BIBLIOGRAFÍA

- Chica, J., Salmerón, R.; 2019 - *Econometría II*
- RedHistoria; 2018 - Los orígenes de los juegos de azar (<https://redhistoria.com/los-origenes-de-los-juegos-de-azar/#:~:text=Los%20primeros%20juegos%20de%20azar,diferentes%2C%20siendo%20el%20precursor%20del>)

- Casi-casino; 2019 - Breve historia de los casinos (<https://www.casi-casino.com/historia-de-los-casinos/#:-:~:text=El%20primer%20casino%20de%20la,muy%20parecido%20al%20Black%2DJack.>)
- Dostoievski, F.; 1867 - El jugador ([https://es.wikisource.org/wiki/El\\_jugador](https://es.wikisource.org/wiki/El_jugador))
- Palumberi, E., Mannino, G.; 2008 - Ludopatía. Un estudio comparativo realizado en Italia y España (<http://papelesdesociedad.info/IMG/pdf/67811462007.pdf>)
- Periódico ABC; 2019 - *España cuenta con la tasa más alta de Europa de ludópatas entre 14 y 21 años* ([https://www.abc.es/familia/padres-hijos/abci-espana-cuenta-tasa-mas-alta-europa-ludopatas-entre-14-y-21-anos-201909110151\\_noticia.html](https://www.abc.es/familia/padres-hijos/abci-espana-cuenta-tasa-mas-alta-europa-ludopatas-entre-14-y-21-anos-201909110151_noticia.html))
- La Vanguardia; 2019 - El perfil del ludópata en 2019 es el de hombre de 36 años, con educación secundaria y trabajo fijo (<https://www.lavanguardia.com/local/galicia/20190922/47551326478/el-perfil-del-ludopata-en-2019-es-el-de-hombre-de-36-anos-con-educacion-secundaria-y-trabajo-fijo.html>)
- Ibáñez, A., Saiz, J.; 2001 - Epidemiología de la ludopatía (Juego Patológico) ([https://psiquiatria.com/article.php?ar=trastorno\\_control\\_impulsos&wurl=epidemiologia-de-la-ludopatia-juego-patologico](https://psiquiatria.com/article.php?ar=trastorno_control_impulsos&wurl=epidemiologia-de-la-ludopatia-juego-patologico))

## BIBLIOGRAFÍA ESPECÍFICA

- Ministerio de Consumo; 2015 - Prevalencia, comportamiento y características de los usuarios de juegos de azar en España (<https://ordenacionjuego.es/es/estudio-prevalencia>)
- Ibáñez, A., Saiz, J.; 2001 - Epidemiología de la ludopatía (Juego Patológico) ([https://psiquiatria.com/article.php?ar=trastorno\\_control\\_impulsos&wurl=epidemiologia-de-la-ludopatia-juego-patologico](https://psiquiatria.com/article.php?ar=trastorno_control_impulsos&wurl=epidemiologia-de-la-ludopatia-juego-patologico))