

Análisis de factores de riesgo en miocardiopatía arritmogénica

Carla Martín Pérez

Modelización de datos categóricos. Grado en Estadística. UGR



**UNIVERSIDAD
DE GRANADA**

Fecha: 29 de mayo de 2025

Índice

1. Introducción	2
2. Metodología y resultados	4
2.1. Modelo log-lineal	4
2.1.1. Fórmula del modelo log-lineal	4
2.1.2. Grafo	5
2.1.3. Residuos	6
2.1.4. Estimaciones de los parámetros del modelo log-lineal ajustado	6
2.2. Modelo logit	8
2.2.1. Fórmula del modelo logit	8
2.2.2. Modelo equivalente	8
3. Conclusiones	11
4. Anexo	12
4.1. Código	15
5. Bibliografía	19

1. Introducción

En este trabajo se estudian los datos de pacientes reales, los cuales han estado bajo seguimiento entre los meses de febrero y junio de 2023 y se ha podido recoger información sobre distintas pruebas médicas y síntomas. Se excluyeron aquellos con contraindicaciones absolutas para la ergometría, como el embarazo, limitaciones físicas para el ejercicio, valvulopatías graves, antecedentes recientes de síndrome coronario agudo o episodios arrítmicos severos, así como enfermedad coronaria no tratada.

Estos pacientes tienen la enfermedad miocardiopatía arritmogénica (MCA) es una enfermedad hereditaria del músculo cardíaco. Exactamente se encuentran pacientes con la enfermedad miocardiopatía arritmogénica con afectación del ventrículo izquierdo, MCAVI, y un grupo control de individuos sanos, los cuales han sido sometidos a pruebas de ergometría.

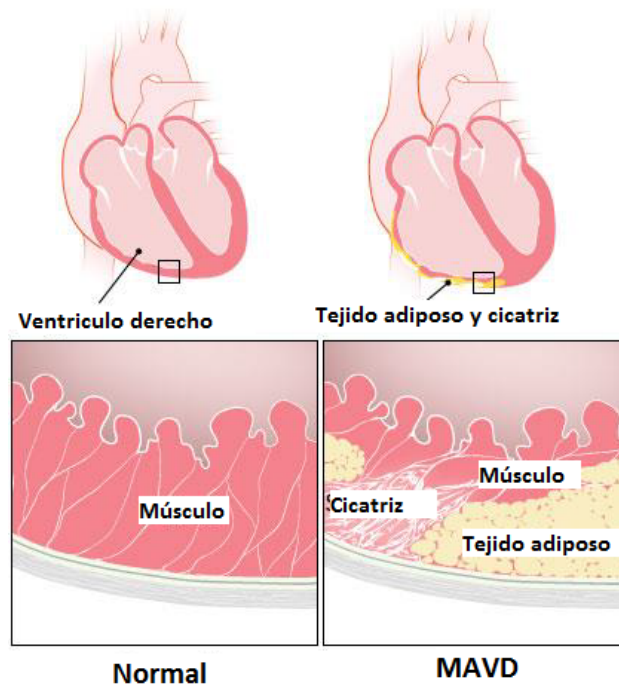


Figura 1: Comparación de un corazón sano y un corazón con MCAVD

Algunos estudios que han analizado la asociación de la miocardiopatía arritmogénica (MCA) con factores como la edad, sexo, fumar y presencia de EVBRD:

- **Sexo.** Un estudio publicado en *EP Europace* estudió las diferencias entre el sexo y la progresión de la enfermedad y el riesgo arrítmico en pacientes con MCA. Se encontró que el sexo masculino se asociaba con una enfermedad más severa. Sin embargo, al ajustar por la exposición al ejercicio, el sexo masculino ya no fue un marcador de riesgo para arritmias ventriculares, sugiriendo que la mayor exposición al ejercicio en hombres podría haber sido un factor de confusión en estudios previos. [1]

- **Fumar.** El hábito de fumar es un factor de riesgo para enfermedades cardíacas conocido y en un estudio publicado por la *Fundación Española del Corazón* se ha concluido que aunque este factor no tenga una relación directa con la MCA aumenta el riesgo de desarrollar enfermedades coronarias y puede reducir el flujo sanguíneo. [2]
- **Edad.** Un estudio recogido en la Revista Española de Cardiología obtuvo como conclusión que la MCA es una de las mayores causas de muerte súbita en adolescentes y adultos jóvenes. [3]
- **EV-BRD.** La presencia de EV-BRD es uno de los criterios menores para el diagnóstico de esta enfermedad, criterio recogido en los *Criterios de Padua*. [4]

Concretamente, este trabajo tiene como objetivo analizar la dependencia y asociación de presentar EV-BRD con tener o no la enfermedad MCAVI junto a otros factores de riesgo como son la edad, ser o no fumador y el sexo .

Para la realización de esta actividad se ha elegido las siguientes variables:

1. **Grupos.** Esta es una variable categórica nominal, encontrándonos en ella dos grupos, uno enfermo y otro sano, con las siguiente categorías: *0:MCAVI/1:Sanos*
2. **Fumar.** En esta variable categórica dicotómica se indica si la persona es fumadora o no. (*0: No/1: Sí*)
3. **Presencia de extrasístoles con bloqueo de rama derecha (EV-BRD).** Esta es una variable categórica dicotómica, donde se observa si un individuo ha presentado extrasístoles con bloqueo de rama derecha en una prueba de esfuerzo. (*0: No/1: Sí*)
4. **Sexo.** Esta es una variable categórica nominal, ya que tiene las siguiente categorías: *0: Mujer/1: Hombre*.
5. **Edad,** como variable estrato. Esta variable está dividida en 7 categorías las cuales son rangos de edad, *1: 0-14, 2: 15-20, 3: 21-30, 4: 31-40, 5: 41-50, 6: 51-60, 7: >61*.

2. Metodología y resultados

2.1. Modelo log-lineal

Para comenzar se contrasta si un modelo que incluye determinados parámetros de asociación se adapta mejor a los datos que otro modelo que no los incorpora siendo el más parsimonioso, con el objetivo de obtener el modelo log-lineal.

Para obtener este modelo se usará el método *backward*, método que parte del modelo más complejo, es decir, el que incluye todos los efectos (modelo saturado), y va eliminando efectos hacia atrás. Por lo tanto, este método en cada etapa contrasta cada modelo resultante de la eliminación de términos de la etapa justamente anterior que, basándose en el p-valor, compara el p-valor de cada parámetro, eliminando el mayor p-valor entre los que son mayores que el nivel de significación. Este método termina cuando cada uno de los parámetros tienen un p-valor menor que el nivel de significación, resultando el modelo más parsimonioso y que se ajusta bien a los datos.

	X^2	df	p-valor
Likelihood Ratio	50.983	90	0.999

Tabla 1: Contraste de bondad de ajuste

Tras aplicar este método, se ha obtenido que el modelo más parsimonioso y que mejor se ajusta a los datos es:

(GB, FB, GE, FS, BS, FBS)

Donde:

- **G**: efecto de la variable GROUP.
- **F**: efecto de la variable Fumar.
- **B**: efecto de la variable EVBRD.
- **E**: efecto de la variable Edad.
- **S**: efecto de la variable Sexo.

A partir de los resultados obtenidos en la tabla, no hay evidencias estadísticas significativas para rechazar la hipótesis nula ya que el p-valor > 0.05 , por lo que se ha obtenido el objetivo de este apartado.

2.1.1. Fórmula del modelo log-lineal

$$\ln(m_{ijklm}) = \mu + \lambda_i^G + \lambda_j^F + \lambda_k^B + \lambda_l^E + \lambda_m^S + \lambda_{ik}^{GB} + \lambda_{jk}^{FB} + \lambda_{il}^{GE} + \lambda_{jm}^{FS} + \lambda_{km}^{BS} + \lambda_{jkm}^{FBS},$$

$i, j, k, l = 1, 2, m = 2, 3, \dots, 7.$

2.1.2. Grafo

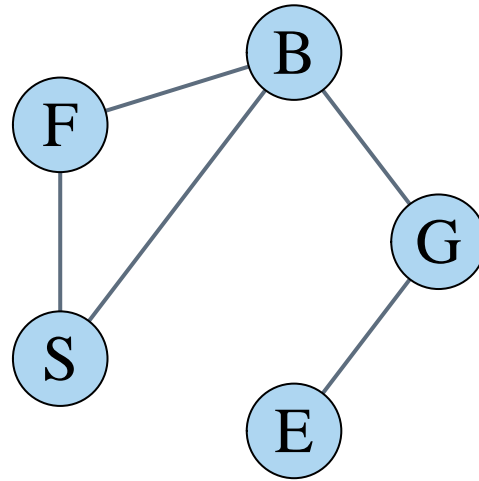


Figura 2: Grafo del modelo log-lineal

En este grafo queda representada cada una de las relaciones de las variables, donde se observa que:

- La variable **Edad** es condicionalmente independiente de Presencia de extrasístole, Fumar y Sexo dada la variable GROUP.
- La variable **Group** es condicionalmente independiente de Sexo y Fumar dada la Presencia de extrasístole.

Aparte de analizar la independencia, se puede analizar la colapsabilidad:

- La tabla es colapsable en **Edad**, y la asociación de GROUP, Sexo, Fumar y EVBRD se podría estudiar en la tabla marginal $n_{i,j,k,\bullet,m}$.
- La tabla es colapsable en **Edad** y **GROUP**, y la asociación de Sexo, Fumar y EVBRD se podría estudiar en la tabla marginal tridimensional $n_{\bullet,j,k,\bullet,m}$.
- La tabla es colapsable en **Sexo**, **Fumar** y **EVBRD**, y la asociación de Edad y GROUP se podría estudiar en la tabla bidimensional $n_{i,\bullet,\bullet,l,m}$.

2.1.3. Residuos

Tras obtener el modelo y saber que se ajusta bien a los datos, se realizará un análisis de los residuos para conocer como de bien se ajusta en cada uno de los posibles cruces de categorías.

Hipótesis:

$$\begin{cases} H_0 : r_i = 0 \\ H_1 : r_i \neq 0 \end{cases}$$

En el apartado de *Anexo* se encuentra la *Tabla 2: Análisis de los residuos*, la cual vienen representados los residuos de cada una de las combinaciones posibles de las cinco variables consideradas en el modelo. Tras revisar dichos residuos se observa que únicamente una combinación presenta un residuo significativamente distinto de cero (en valor absoluto superior a 1.96), lo cual sugiere un posible desajuste puntual del modelo.

Esta celda corresponde a la combinación Sano, no fumador, presenta extrasístoles, hombre y mayor de 61 años con un solo individuo por lo que el porcentaje de desajuste es menor de 0.01.

En conjunto, el análisis de los residuo no revela patrones sistemáticos de desajuste, por lo que se considera que el modelo ajusta adecuadamente a los datos.

2.1.4. Esimaciones de los parámetros del modelo log-lineal ajustado

Efecto	λ	Error típico	Z	IC 95 %	exp λ
Constante	-5.247	—	—	—	0.005
GROUP: MCAVI (G)*	3.688	1.017	3.626	(1.696, 5.681)	40
Fumar: No (F)*	2.485	1.041	2.387	(0.445, 4.525)	12
EVBRD: No (B)	2.028	1.583	1.281	(-1.075, 5.131)	7.6
Edad: 0-14 (E)	-18.082	3620.127	-0.005	(-7113.401, 7077.238)	0
Edad: 15-20 (E)	-18.082	3620.127	-0.005	(-7113.401, 7077.238)	0
Edad: 21-30 (E)	-0.693	1.225	-0.566	(-3.094, 1.707)	0.5
Edad: 31-40 (E)	1.099	0.816	1.347	(-0.502, 2.698)	3
Edad: 41-50 (E)	-0.693	1.225	-0.566	(-3.094, 1.707)	0.5
Edad: 51-60 (E)*	1.609	0.775	2.076	(0.091, 3.127)	5
Sexo: Mujer	-16.68	2540.397	-0,007	(-4995.768, 4962.407)	0
GB: MCAVI-No*	-2.485	0.788	-3.154	(-4.028, -0.941)	0.083
FB: No-No	0.734	1.457	0.504	(-2.122, 3.589)	2.083
GE: MCAVI-[0-14]	-2.228	5521.395	0.0004	(-10823.96, 10819.51)	0.108

Efecto	λ	Error típico	Z	IC 95 %	exp λ
GE:	16.459	3620.128	0.005	(-7078.778, 7111.861)	15267410
MCAVI-[15-20]					
GE:	1.782	1.309	1.361	(-2.566, 2.566)	1
MCAVI-[21-30]					
GE:	-1.792	0.938	-1.91	(-3.631, 0.048)	0.167
MCAVI-[31-40]					
GE:	0.762	1.279	0.596	(-1.746, 3.271)	2.143
MCAVI-[41-50]					
GE:	-1.415	0.855	-1.655	(-3.09, 0.259)	0.243
MCAVI-[51-60]					
FS: No-Mujer	17.414	2540.397	0.007	(-4961.673, 4996.502)	36547400
BS: No-Mujer	18.066	2540.398	0.007	(-4961.022, 4997.154)	70171010
FBS:	-19.311	2540.398	-0.008	(-4998.399, 4959.777)	0
No-No-Mujer					

Tabla 3: Estimación de los parámetros

En esta tabla se han obtenido las estimaciones de los parámetros junto a su exponencial, intervalos de confianza error típico y Z-valor, siendo algunos de ellos significativos ya que el valor de **Z** en valor absoluto es mayor que 1.96. Aún así, la mayoría de ellos tienen un Z-valor menor de 1.96, los cuales no son significativos. Los que han resultado significativos estarán señalados con un *.

Una breve interpretación de lo obtenido en la tabla:

- **Parámetro constante.** Indica el número esperado de pacientes sanos, que fuman, que presentan EVBRD, con una edad mayor o igual de 61 y varones es 0.005.
- **Efectos principales.** Se refiere a las exponenciales de los efectos principales de una variable, que son las ventajas a favor de la categoría asociada en lugar de la última para los individuos que están en las últimas categorías del resto de variables.
 - **Variable Group.** La ventaja de ser diagnosticado con MCAVI para pacientes varones, que presenten EVBRD, fuman y tengan una edad de 61 años o mayor es de 40 frente a ser diagnosticado como sano.
 - **Variable Fumar.** La ventaja de no fumar para pacientes varones, sanos, que presenten EVBRD y tengan una edad de 61 años o mayor es de 12 frente a fumar.
 - **Variable Edad.** La ventaja de tener entre 51 y 60 años para pacientes varones, sanos, que presenten EVBRD y sean fumadores es de 5 frente a tener 61 años o más.
- **Interacciones.** Se utilizarán las exponenciales de las interacciones entre cada par de variables como los cocientes de ventajas ordinarios para los individuos que pertenecen a la última categoría de la variable no involucrada.

- La ventaja de ser diagnosticado de MCAVI y no presentar EVBRD es de 0.083 para pacientes varones, fumadores y con 61 años o más.

Después de haber obtenido el modelo log-lineal que mejor se ajusta a los datos, se puede ajustar un modelo logit.

2.2. Modelo logit

Teniendo en cuenta el modelo anterior, donde se seleccionó el modelo más parsimonioso que se ajusta bien a la tabla, que fue el modelo {GB, FB, GE, FS, BS, FBS}, no se diferenció entre variables explicativas y de respuesta, en este modelo se hará. Se ajustará un modelo logit que trate ‘Presencia de EVBRD’ como variable respuesta y las variables ‘Group’ ‘Fumar’, ‘Sexo’ y ‘Edad’ como variables explicativas.

Expresión log-lineal: $\ln(m_{ijklm}) = \mu + \lambda_i^G + \lambda_j^F + \lambda_k^B + \lambda_l^E + \lambda_m^S + \lambda_{ik}^{GB} + \lambda_{jk}^{FB} + \lambda_{il}^{GE} + \lambda_{jm}^{FS} + \lambda_{km}^{BS} + \lambda_{jkm}^{FBS}$,
 $i, j, k, l = 1, 2, m = 2, 3, \dots, 7$.

2.2.1. Fórmula del modelo logit

Este modelo log-lineal implica este **modelo logit**:

$$l_{ijklm} = \ln\left(\frac{m_{ij1lm}}{m_{ij2lm}}\right) =$$

$$(\mu + \lambda_i^G + \lambda_j^F + \lambda_1^B + \lambda_l^E + \lambda_m^S + \lambda_{i1}^{GB} + \lambda_{j1}^{FB} + \lambda_{il}^{GE} + \lambda_{jm}^{FS} + \lambda_{1m}^{BS} + \lambda_{j1m}^{FBS}) -$$

$$(\mu + \lambda_i^G + \lambda_j^F + \lambda_2^B + \lambda_l^E + \lambda_m^S + \lambda_{i2}^{GB} + \lambda_{j2}^{FB} + \lambda_{il}^{GE} + \lambda_{jm}^{FS} + \lambda_{2m}^{BS} + \lambda_{j2m}^{FBS}) =$$

$$(\lambda_1^B - \lambda_2^B) + (\lambda_{i1}^{GB} - \lambda_{i2}^{GB}) + (\lambda_{j1}^{FB} - \lambda_{j2}^{FB}) + (\lambda_{1m}^{BS} - \lambda_{2m}^{BS}) + (\lambda_{j1m}^{FBS} - \lambda_{j2m}^{FBS}) =$$

$$\alpha + \tau_i^G + \tau_j^F + \tau_m^S + \tau_{jm}^{FS}$$

2.2.2. Modelo equivalente

Se comprueba si este modelo logit es equivalente al modelo log-lineal seleccionado {GB, FB, GE, FS, BS, FBS}:

1. *Contiene el efecto principal de la variable respuesta.* Como aparece el término λ_k^B , se puede verificar esta hipótesis.
2. *Contiene las interacciones entre la variable respuesta y cada una de las variables explicativas que aparecen en el modelo logit.* En este caso se cumple, ya que el modelo contiene el término

λ_{ik}^{GB} asociado a τ_i^G , el término λ_{jk}^{FB} asociado a τ_j^F , el término λ_{km}^{BS} asociado a τ_m^S , y el término λ_{jkm}^{FBS} asociado a τ_{jm}^{FS} .

3. *Contiene la interacción de mayor orden entre todas las variables explicativas de la tabla.* Esta condición no se verifica puesto que el modelo log-lineal no contiene la interacción entre Group, Fumar, Sexo y Edad.

Por tanto, el modelo {GB, FB, GE, FS, BS, FBS} no es equivalente, y se considerará el modelo equivalente {GB, FB, BS, FBS, GFES}.

	X^2	df	p-valor
Likelihood Ratio	22.434	51	0.999

Las estimaciones pertenecientes a este modelo son:

$$\hat{\alpha} = (\lambda_1^B - \lambda_2^B), \hat{\tau}_i^G = (\lambda_{i1}^{GB} - \lambda_{i2}^{GB})$$

,

$$\hat{\tau}_j^F = (\lambda_{j1}^{FB} - \lambda_{j2}^{FB}), \hat{\tau}_m^S = (\lambda_{1m}^{BS} - \lambda_{2m}^{BS}), \hat{\tau}_{jm}^{FS} = (\lambda_{j1m}^{FBS} - \lambda_{j2m}^{FBS})$$

La estimación de los parámetros por máxima verosimilitud de este modelo log-lineal se ha realizado mediante el método marginal. En la siguiente tabla se recogen los parámetros estimados del modelo log-lineal equivalente {GB, FB, BS, FBS, GFES}, junto con el valor Z correspondiente.

Parámetros	M. Parcial (Z.valor)
No = $\lambda_1^B = \hat{\alpha}$	2.028 (1.281)
MCAVI-No = $\lambda_{11}^{GB} = \hat{\tau}_1^G$	-2.485 (-3.154)
No-No = $\lambda_{11}^{FB} = \hat{\tau}_1^F$	0.734 (0.504)
No-Mujer = $\lambda_{11}^{BS} = \hat{\tau}_1^S$	18.066 (0.007)
No-No-Mujer = $\lambda_{111}^{FBS} = \hat{\tau}_{11}^{FS}$	-19.311 (-0.008)

La interpretación de estas estimaciones son:

- $\hat{\alpha}$. La ventaja a favor de no presentar EVBRD sería $e^{2,028} = 7,6$, pero al ser no significativo no existen diferencias en la probabilidad de presentar EVBRD si el paciente es sano, fumador, hombre y mayor de 61 años.
- El efecto del grupo sobre la presencia de extrasístoles, $\hat{\tau}_1^G$, viene dado por el cocientes de ventajas

$$\hat{\theta}^{GB} = \exp(\hat{\tau}_1^G - \hat{\tau}_2^G) = e^{-2,485} = 0,083 \rightarrow \frac{1}{0,083} = 12,05$$

Esto es, la probabilidad de presentar EVBRD es 12.05 veces mayor para los que son portadores de la enfermedad MCAVI que para los que son sanos.

- El efecto de fumar sobre la presencia de extrasístoles, $\hat{\tau}_1^F$, viene dado por el cocientes de ventajas

$$\hat{\theta}^{FB} = \exp(\hat{\tau}_1^F + \hat{\tau}_{11}^{FS}) = e^{0,734-19,311} = 8,552861e-09 \longrightarrow \frac{1}{8,552861e-09} = 116919941$$

Al no ser significativo, no hay diferencias en la probabilidad de presentar EVBRD si el pacientes es fumador o no.

- El efecto del sexo sobre la presencia de extrasístoles, $\hat{\tau}_1^S$, viene dado por el cocientes de ventajas

$$\hat{\theta}^{BS} = \exp(\hat{\tau}_1^S + \hat{\tau}_{11}^{FS}) = e^{18,066-19,311} = 0,288 \longrightarrow \frac{1}{0,288} = 3,47$$

Esto es, la probabilidad de presentar EVBRD es 3.47 veces mayor para las mujeres no fumadoras que para los hombres fumadores.

La probabilidad de presentar EVBRD es $\exp(\hat{\tau}_1) = 70139733 \longrightarrow \frac{1}{70139733} = 0$, es decir, es nula para los casos en los que hay hombres no fumadores.

3. Conclusiones

El objetivo principal de este trabajo ha sido analizar la asociación entre la presencia de extrasístoles ventriculares con morfología de bloqueo de rama derecha (EVBRD) y la miocardiopatía arritmogénica con afectación del ventrículo izquierdo (MCAVI), considerando además otros factores de riesgo como el sexo, el hábito de fumar y la edad.

Mediante el ajuste de un modelo log-lineal y un modelo logit, se ha comprobado que el modelo log-lineal seleccionado {GB, FB, BS, FBS, GFES} proporciona un buen ajuste a los datos y sin evidencia de desajuste sistemáticos en los residuos, excepto en una combinación puntual.

Las estimaciones del modelo logit permiten interpretar las asociaciones entre variables. Entre las conclusiones destacadas se encuentran:

- La presencia de EVBRD está fuertemente asociada a padecer MCAVI, ya que los pacientes con la enfermedad tienen una probabilidad estimada 12 veces mayor de presentar EVBRD en comparación con los individuos sanos.
- El efecto del sexo sobre la presencia de EVBRD muestra que, dentro del grupo de mujeres no fumadoras, la probabilidad es 3.47 veces mayor que en hombre fumadores, esto puede deberse a que las personas afectadas por la enfermedad no tienen el hábito de fumar. Sin embargo, estos efectos no se han visto que sean significativos estadísticamente.
- Al efecto de fumar le ocurre lo mismo que al efecto del sexo sobre la presencia de EVBRD, ya que no existe evidencia significativa para afirmar que fumar esté asociado con presentar EVBRD.
- Respecto al efecto de la edad, se ha obtenido que es independiente de la presencia de EVBRD, por lo que no existe una asociación entre ellos.

Finalmente, el estudio permite concluir que la miocardiopatía arritmogénica con afectación del ventrículo izquierdo tiene una relación estadísticamente significativa con la aparición de EVBRD.

Además, señalar las limitaciones del estudio debido al tamaño reducido de la muestra, ya que se tienen pocas observaciones, por lo que es más difícil detectar asociaciones entre las variables, lo que puede dar lugar a errores tipo II, no detectar un efecto que sí existe. Este tamaño de la muestra también ha provocado que existan celdas con 0 o 1 observaciones, dificultando la estimación de parámetros e interacciones.

4. Anexo

	GROUP	Fumar	EVBRD	Sexo	Edad	Freq	residuos_estandar
1	MCAVI	No	No	Mujer	0-14	0	-8.399652e-05
2	Sanos	No	No	Mujer	0-14	0	-1.551156e-04
3	MCAVI	Sí	No	Mujer	0-14	0	-4.094551e-05
4	Sanos	Sí	No	Mujer	0-14	0	-6.987068e-05
5	MCAVI	No	Sí	Mujer	0-14	0	-1.596919e-04
6	Sanos	No	Sí	Mujer	0-14	0	-6.287815e-05
7	MCAVI	Sí	Sí	Mujer	0-14	0	-2.107342e-08
8	Sanos	Sí	Sí	Mujer	0-14	0	-2.107342e-08
9	MCAVI	No	No	Hombre	0-14	0	-1.150170e-04
10	Sanos	No	No	Hombre	0-14	0	-2.369430e-04
11	MCAVI	Sí	No	Hombre	0-14	0	-2.017531e-05
12	Sanos	Sí	No	Hombre	0-14	0	-3.384900e-05
13	MCAVI	No	Sí	Hombre	0-14	0	-9.654489e-05
14	Sanos	No	Sí	Hombre	0-14	0	-4.278688e-05
15	MCAVI	Sí	Sí	Hombre	0-14	0	-2.542238e-05
16	Sanos	Sí	Sí	Hombre	0-14	0	-1.217094e-05
17	MCAVI	No	No	Mujer	15-20	1	8.157985e-01
18	Sanos	No	No	Mujer	15-20	0	-1.551156e-04
19	MCAVI	Sí	No	Mujer	15-20	0	-4.948396e-01
20	Sanos	Sí	No	Mujer	15-20	0	-6.987068e-05
21	MCAVI	No	Sí	Mujer	15-20	1	-1.591030e-01
22	Sanos	No	Sí	Mujer	15-20	0	-6.287815e-05
23	MCAVI	Sí	Sí	Mujer	15-20	0	-7.339051e-05
24	Sanos	Sí	Sí	Mujer	15-20	0	-2.107342e-08
25	MCAVI	No	No	Hombre	15-20	1	3.718713e-01
26	Sanos	No	No	Hombre	15-20	0	-2.369430e-04
27	MCAVI	Sí	No	Hombre	15-20	0	-2.437295e-01
28	Sanos	Sí	No	Hombre	15-20	0	-3.384900e-05
29	MCAVI	No	Sí	Hombre	15-20	0	-1.176484e+00
30	Sanos	No	Sí	Hombre	15-20	0	-4.278688e-05
31	MCAVI	Sí	Sí	Hombre	15-20	0	-3.097316e-01
32	Sanos	Sí	Sí	Hombre	15-20	0	-1.217094e-05
33	MCAVI	No	No	Mujer	21-30	0	-1.589572e+00
34	Sanos	No	No	Mujer	21-30	1	1.213035e+00
35	MCAVI	Sí	No	Mujer	21-30	0	-7.717234e-01
36	Sanos	Sí	No	Mujer	21-30	0	-4.213250e-01
37	MCAVI	No	Sí	Mujer	21-30	1	-1.544467e+00

38	Sanos	No	Sí	Mujer	21-30	0	-3.815750e-01
39	MCAVI	Sí	Sí	Mujer	21-30	0	-1.158128e-04
40	Sanos	Sí	Sí	Mujer	21-30	0	-1.734389e-05
41	MCAVI	No	No	Hombre	21-30	4	1.831134e+00
42	Sanos	No	No	Hombre	21-30	0	-1.431007e+00
43	MCAVI	Sí	No	Hombre	21-30	0	-3.798932e-01
44	Sanos	Sí	No	Hombre	21-30	0	-2.040887e-01
45	MCAVI	No	Sí	Hombre	21-30	2	7.053675e-01
46	Sanos	No	Sí	Hombre	21-30	0	-2.575146e-01
47	MCAVI	Sí	Sí	Hombre	21-30	0	-4.887783e-01
48	Sanos	Sí	Sí	Hombre	21-30	0	-7.278010e-02
49	MCAVI	No	No	Mujer	31-40	2	9.879598e-01
50	Sanos	No	No	Mujer	31-40	3	1.044073e+00
51	MCAVI	Sí	No	Mujer	31-40	0	-7.717234e-01
52	Sanos	Sí	No	Mujer	31-40	0	-1.090009e+00
53	MCAVI	No	Sí	Mujer	31-40	1	-1.544467e+00
54	Sanos	No	Sí	Mujer	31-40	0	-1.025174e+00
55	MCAVI	Sí	Sí	Mujer	31-40	0	-1.158128e-04
56	Sanos	Sí	Sí	Mujer	31-40	0	-4.276672e-05
57	MCAVI	No	No	Hombre	31-40	1	-6.687118e-01
58	Sanos	No	No	Hombre	31-40	3	3.213667e-08
59	MCAVI	Sí	No	Hombre	31-40	0	-3.798932e-01
60	Sanos	Sí	No	Hombre	31-40	0	-5.276573e-01
61	MCAVI	No	Sí	Hombre	31-40	2	7.053675e-01
62	Sanos	No	Sí	Hombre	31-40	0	-6.590873e-01
63	MCAVI	Sí	Sí	Hombre	31-40	1	1.754658e+00
64	Sanos	Sí	Sí	Hombre	31-40	0	-1.799787e-01
65	MCAVI	No	No	Mujer	41-50	3	6.519337e-01
66	Sanos	No	No	Mujer	41-50	0	-9.359019e-01
67	MCAVI	Sí	No	Mujer	41-50	1	5.654697e-01
68	Sanos	Sí	No	Mujer	41-50	0	-4.213250e-01
69	MCAVI	No	Sí	Mujer	41-50	9	1.878633e+00
70	Sanos	No	Sí	Mujer	41-50	0	-3.815750e-01
71	MCAVI	Sí	Sí	Mujer	41-50	0	-1.822251e-04
72	Sanos	Sí	Sí	Mujer	41-50	0	-1.734389e-05
73	MCAVI	No	No	Hombre	41-50	1	-2.041955e+00
74	Sanos	No	No	Hombre	41-50	1	8.894079e-01
75	MCAVI	Sí	No	Hombre	41-50	0	-5.805330e-01
76	Sanos	Sí	No	Hombre	41-50	0	-2.040887e-01
77	MCAVI	No	Sí	Hombre	41-50	1	-1.498837e+00

78	Sanos	No	Sí	Hombre	41-50	0	-2.575146e-01
79	MCAVI	Sí	Sí	Hombre	41-50	0	-7.691082e-01
80	Sanos	Sí	Sí	Hombre	41-50	0	-7.278010e-02
81	MCAVI	No	No	Mujer	51-60	0	-2.638290e+00
82	Sanos	No	No	Mujer	51-60	3	0.000000e+00
83	MCAVI	Sí	No	Mujer	51-60	1	4.527733e-01
84	Sanos	Sí	No	Mujer	51-60	1	2.512345e-01
85	MCAVI	No	Sí	Mujer	51-60	8	9.021707e-01
86	Sanos	No	Sí	Mujer	51-60	1	4.934747e-01
87	MCAVI	Sí	Sí	Mujer	51-60	0	-1.978737e-04
88	Sanos	Sí	Sí	Mujer	51-60	0	-5.550924e-05
89	MCAVI	No	No	Hombre	51-60	5	5.871575e-01
90	Sanos	No	No	Hombre	51-60	5	0.000000e+00
91	MCAVI	Sí	No	Hombre	51-60	1	1.539944e+00
92	Sanos	Sí	No	Hombre	51-60	0	-7.145896e-01
93	MCAVI	No	Sí	Hombre	51-60	2	-8.352558e-01
94	Sanos	No	Sí	Hombre	51-60	0	-8.839441e-01
95	MCAVI	Sí	Sí	Hombre	51-60	0	-8.351692e-01
96	Sanos	Sí	Sí	Hombre	51-60	0	-2.341583e-01
97	MCAVI	No	No	Mujer	>61	2	0.000000e+00
98	Sanos	No	No	Mujer	>61	0	-1.338299e+00
99	MCAVI	Sí	No	Mujer	>61	0	-1.134231e+00
100	Sanos	Sí	No	Mujer	>61	1	1.499684e+00
101	MCAVI	No	Sí	Mujer	>61	4	-8.221912e-01
102	Sanos	No	Sí	Mujer	>61	0	-5.489736e-01
103	MCAVI	Sí	Sí	Mujer	>61	0	-1.743615e-04
104	Sanos	Sí	Sí	Mujer	>61	0	-2.456039e-05
105	MCAVI	No	No	Hombre	>61	4	4.451204e-01
106	Sanos	No	No	Hombre	>61	0	-2.048366e+00
107	MCAVI	Sí	No	Hombre	>61	0	-5.577274e-01
108	Sanos	Sí	No	Hombre	>61	0	-2.916266e-01
109	MCAVI	No	Sí	Hombre	>61	4	1.062690e+00
110	Sanos	No	Sí	Hombre	>61	1	1.974462e+00
111	MCAVI	Sí	Sí	Hombre	>61	0	-7.359128e-01
112	Sanos	Sí	Sí	Hombre	>61	0	-1.031212e-01

Tabla 2: Análisis de los residuos

4.1. Código

```
knitr::opts_chunk$set(echo=TRUE, message=T, warning=F, eval=TRUE, comment="", results="markup")
# comment evita el que aparezca el símbolo ## en el
library(readxl)
library(ggplot2)
library(tidyverse)
library(scales)

#Datos que vamos a usar
datos <- read_xlsx("basedatos.xlsx", col_names = T)
datos <- as.data.frame(datos)
datos

datos <- datos[, c(2, 6, 94, 5, 4)]
datos <- datos[-c(64:83),]
colnames(datos) <- c("GROUP", "Fumar", "EVBRD", "Sexo", "Edad")
head(datos)

datos$GROUP <- factor(datos$GROUP, levels = c(0, 1), labels = c("MCAVI", "Sanos"))

datos$Fumar <- factor(datos$Fumar, levels = c(0, 1), labels = c("No", "Sí"))

datos$EVBRD <- factor(datos$EVBRD, levels = c(0, 1), labels = c("No", "Sí"))

datos$Sexo <- factor(datos$Sexo, levels = c(0, 1), labels = c("Mujer", "Hombre"))

datos$Edad <- factor(datos$Edad, levels = 1:7, labels = c("0-14", "15-20", "21-30", "31-40", "41-50", "51-60", "61-70"))

datos <- table(datos)
datos

datos <- as.data.frame(datos)
datos

require(MASS)

saturado = glm(Freq ~GROUP*Fumar*EVBRD*Edad*Sexo , family = poisson, data = datos)

names(saturado)

saturado$param
```



```
summary(saturado)
```

```
modelo_final <- step(saturado, direction="backward", test="Chisq")
modelo_final
```

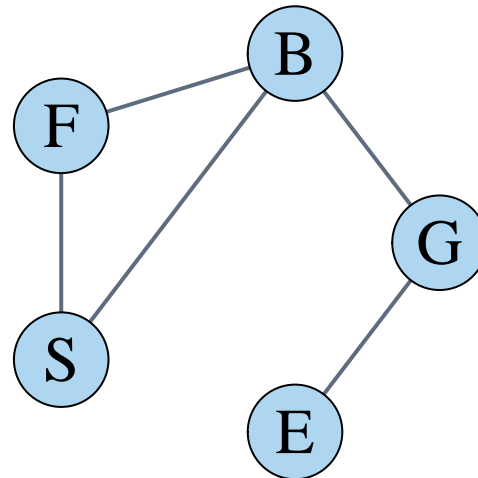
```
summary(loglm(formula = Freq ~ GROUP + Fumar + EVBRD + Edad + Sexo + GROUP:EVBRD +
  Fumar:EVBRD + GROUP:Edad + Fumar:Sexo + EVBRD:Sexo + Fumar:EVBRD:Sexo, data = datos))$test
```

```
library(igraph)
```

```
grafo <- graph_from_literal(
  G -- B, # GB
  F -- B,
  F -- S,
  B -- S,
  G -- E
)
```

```
set.seed(123)
```

```
plot(grafo,
  layout = layout_in_circle,
  vertex.color = "#AED6F1",
  vertex.size = 50,
  vertex.label.cex = 2,
  vertex.label.color = "black",
  edge.width = 2,
  edge.color = "#5D6D7E")
```



```

contrasts(datos$GROUP) = contr.treatment(2, 2)
contrasts(datos$Fumar) = contr.treatment(2, 2)
contrasts(datos$EVBRD) = contr.treatment(2, 2)
contrasts(datos$Sexo) = contr.treatment(2, 2)
contrasts(datos$Edad) = contr.treatment(7, 7)

```

```

modelo.glm = glm(Freq ~ GROUP + Fumar + EVBRD + Edad + Sexo + GROUP:EVBRD +
  Fumar:EVBRD + GROUP:Edad + Fumar:Sexo + EVBRD:Sexo + Fumar:EVBRD:Sexo, family=poisson, data=

```

```

residuos_estandar = rstandard(modelo.glm)
cbind(datos, residuos_estandar)

```

```

#--> Estimamos los parametros del modelo (las ultimas categorias no las muestra).
modeloglm <- glm(formula = Freq ~ GROUP + Fumar + EVBRD + Edad + Sexo + GROUP:EVBRD +
  Fumar:EVBRD + GROUP:Edad + Fumar:Sexo + EVBRD:Sexo + Fumar:EVBRD:Sexo, family=poisson, data=
parametros_glm=modeloglm$coefficients
parametros_glm

```

```

#--> Calculamos la desviacion tipica de los parametros.

```

```

error_estandar_param=summary(modeloglm)$coefficients[,2]
error_estandar_param

#--> Mostramos la tabla 5 del informe

ic = function(parametro, sigma, conf.level=0.95) {
  inf=parametro-qnrm(0.5*(1+conf.level))*sigma
  sup=parametro+qnrm(0.5*(1+conf.level))*sigma
  return(list(inf,sup))
}

tabla_resultados = function(est_parametros=parametros_glm,error_est=error_estandar_param){
  estimaciones=vector()
  exp_estimaciones=vector()
  IC_estim=matrix(data=NA,nrow = length(est_parametros), ncol = 2)
  IC_exp_estim=matrix(data=NA,nrow = length(est_parametros), ncol = 2)

  for(i in 1:length(est_parametros)){
    estimaciones[i]=est_parametros[i]
    exp_estimaciones[i]=exp(est_parametros[i])
    for(j in 1:2){
      IC_estim[i,j]=ic(est_parametros[i],error_est[i],0.95)[[j]]
      IC_exp_estim[i,j]=exp(IC_estim[i,j])
    }
  }

  return(data.frame(names(est_parametros),estimaciones,
exp_estimaciones, IC_inf_est=IC_estim[,1],IC_sup_est=IC_estim[,2], IC_inf_exp_est=IC_exp_estim[,1],
IC_exp_estim[,2]))
}

resultados_completa=tabla_resultados()
resultados_completa

summary(loglm(formula = Freq ~ GROUP + Fumar + EVBRD + Edad + Sexo + GROUP:EVBRD +
  Fumar:EVBRD + EVBRD:Sexo + Fumar:EVBRD:Sexo + GROUP:Fumar:Edad:Sexo, data = datos))$test[1,]

```

5. Bibliografía

- [1] Rootwelt-Norberg, C., Lie, Ø. H., Chivulescu, M., Castrini, A. I., Sarvari, S. I., Lyseggen, E., ... & Haugaa, K. H. (2021). Sex differences in disease progression and arrhythmic risk in patients with arrhythmogenic cardiomyopathy. *EP Europace*, 23(7), 1084-1091.
- [2] Fundación Española del Corazón. (s.f.). Vivir con miocardiopatía arritmogénica. [Cardiopatías familiares. URL](#)
- [3] Aguilera, B., Mier, M. P. S., & Morentin, B. (1999). Miocardiopatía arritmogénica como causa de muerte súbita en España. Presentación de 21 casos. *Revista Española de Cardiología*, 52(9), 656-662.
- [4] Corrado D, Perazzolo Marra M, Zorzi A, et al. Diagnosis of arrhythmogenic cardiomyopathy: The Padua criteria. *Int J Cardiol*. 2020;319:106-114. [doi:10.1016/j.ijcard.2020.06.005](https://doi.org/10.1016/j.ijcard.2020.06.005)