

Trabajo Final: Análisis Estadístico de las Pinturas de la Cueva de la Pileta

Carlos Mota Romero

2023-05-31

1. Introducción:

El arte, y los distintos medios en los que se expresa, parece ser una parte esencial del ser humano, y nos acompaña desde la Prehistoria. El arte parece ser solo practicado por un tipo de homínido, el *Homo sapiens*, y dependiendo de la zona en la que nos encontremos, la cronología en la que este aparece variará. Pero en este estudio no nos centramos en cualquier tipo de arte, sino en el primero del que tenemos constancia. El a veces denominado como arte rupestre, que englobaría al arte parietal y a las piezas móviles sobre soporte pétreo, acompaña a los sapiens en su entrada en Europa desde el Paleolítico Superior (al menos desde 35.000 BP). Aunque algunas producciones anteriores a esta etapa podrían considerarse como arte, esta cronología es la más ampliamente aceptada para el territorio, ya que es cuando estos trabajos comienzan a expandirse, y a evolucionar con una gran rapidez (Sanchidrián, 2001).

En el arte prehistórico europeo encontramos una serie de convenciones que se repiten a lo largo de todo el espacio, sin importar la distancia que exista entre los distintos yacimientos en los que podemos encontrar estos rastros de actividades humanas. La primera de estas convenciones sería el tipo de pigmento, o mejor dicho el color, con el que se elaboran estas representaciones; siempre encontrando solo 3 posibles iteraciones: las pinturas amarillas, las rojas y las negras; lo cuál es el primer indicativo de una conexión ancestral en todo el territorio, ya que en otras zonas podemos encontrar otras coloraciones de pigmento, como las pinturas azules americanas. La segunda de estas convenciones hace referencia a las figuras que se plasman sobre el soporte, que, aunque suelen ser en un inicio motivos naturales, se restringen a una ínfima parte de todos los posibles; es decir, en el arte prehistórico europeo del paleolítico superior (35.000 BP) los motivos más plasmados son los zoomorfos, con una clara y constante repetición de tan solo 4 de ellos: la cabra, el uro, el caballo y el ciervo. Por último, también encontramos que los motivos del arte rupestre europeo en un inicio comienzan siendo ciertamente más naturalistas, y a medida que avanzan en el tiempo son cada vez más esquemáticos, hasta llegar a las marañas postpaleolíticas: trazos informes sobrelapados que no parecen reflejar ningún elemento reconocible para nosotros (Sanchidrián, 2001).

Por las razones que se han expuesto, se puede ver una clara intencionalidad en la repetición de los patrones relacionados con el arte rupestre europeo, por lo que se considera ampliamente que este tipo de representaciones estaban marcadas por una cultura común; una serie de historias que se repetían a lo largo del tiempo y permitían la pervivencia de estos motivos y patrones tan singulares. Es por ello que el arte paleolítico conforma un tema de estudio muy importante para las investigaciones del Paleolítico Superior pues, mediante él, no solo estamos analizando una tipología artística, sino que nos estamos adentrando en uno de los pocos registros culturales que tenemos de esta etapa de la humanidad (Sanchidrián, 2001).

Pero para este, mi trabajo final de la asignatura de Aplicaciones Informáticas en Arqueología y Estadística, he decidido realizar un estudio centrado en las pinturas rupestres halladas en la cueva de “La Pileta” en Benaolán (Málaga). La cueva es conocida por albergar más de 1.200 representaciones en su interior, pero también por su largo periodo de ocupación, que se expande desde el Paleolítico a la Edad Media (Cortés y Simón, 2007).

El objeto de estudio que nos atrae a este lugar es la diversidad y complejidad de los muchos motivos que encontramos en su interior. Los varios autores que han trabajado y realizado investigaciones en La Pileta suelen describir un mínimo de 4 periodos de representaciones, que chocan en estilo y contenidos, y que los hace fácilmente distinguibles. El primer periodo es el que tomaría lugar durante el Solutrense (tercera etapa del Paleolítico Superior 20.000 años), y se relaciona con las pinturas con coloraciones amarillas representando zoomorfos (animales) de manera bastante naturalista. El segundo periodo también dataría del solutrense, pero de un momento más avanzado que el primer grupo descrito; en esta categoría de representaciones, encontramos una gran variedad de pinturas rojas, que también representan animales de manera naturalista. Como tercer bloque tenemos las pinturas del magdaleniense (última etapa del paleolítico superior, de la que se tiene constancia en Pileta, pero de la que no existen dataciones); compuesto por pinturas de coloración negra y motivos también naturalistas. Por último, tenemos las pinturas postpaleolíticas, de coloración oscura y motivos simbólicos muy esquematizados (Cortés y Simón, 2007 y Sanchidrián, 1997).

2. Presentación de las variables y de los objetivos:

La propuesta de secuenciación cronotipológica descrita, ha sido la tradicionalmente usada para el estudio de esta cavidad, pero debemos tener en cuenta que ha sido siempre creada a través de una secuencia cronostilística, que predispone los colores rojo y amarillo y los zoomorfos más realistas a una época más temprana que los motivos más abstractos. Mediante este trabajo, lo que se intenta comprobar es la veracidad y exactitud de esta secuencia, y ver si estadísticamente existe una relación entre los colores de las pinturas y las representaciones en sí mismas; así como la posible relación entre la coloración y la naturalidad o esquematización de los motivos. Para conseguir este propósito, se han elegido aleatoriamente 33 pinturas dentro de la cueva de La Pileta, de las que se observan las siguientes características para cada una de ellas:

1- Figuración: Con dos posibles valores: a. Zoomorfo (Zoo): Si la pintura está representando a un animal reconocible. b. Símbolo (Sim): Si la pintura está representando una idea abstracta no reconocible. La tercera opción debería ser antropomorfo, pero no existen este tipo de representaciones en Pileta, por lo que la opción es obviada.

2- Color: Dependiendo del color de la pintura y con tres posibles valores: a. Amarillo b. Rojo c. Negro Las ocasiones en las que estas pinturas se combinan sobre un mismo dibujo en Pileta son extremadamente escasas, y ninguna de ellas es encontrada en la muestra aleatoria que se toma.

3- Estilo: Variable ordinal que toma los valores del 1 al 5 dependiendo del nivel de esquematización de la pintura muestreada; siendo el 1 una pintura muy naturalista y el 5 una pintura muy esquematizada. Se suele considerar al arte paleolítico como más naturalista y al post-paleolítico como más esquemático.

El objetivo de este trabajo, dadas las variables propuestas, consiste en inferir estadísticamente si existe una relación real entre el color de las pinturas que podemos encontrar en la cueva de La Pileta, con la figuración que representan, tanto en contenido (ya sean zoomorfos o símbolos) y en estilo (más o menos naturalista).

3. Desarrollo de Contenidos

El trabajo estadístico se realiza en el lenguaje de programación R y utilizando el software R-Studio para trabajar el código propuesto. Se expone así, a continuación, el código utilizado para trabajar sobre los objetivos propuestos, así como una explicación del mismo, y un análisis sobre los resultados que se obtienen gracias al mismo.

3.1 Pasos iniciales:

En una primera instancia, debemos descargar y activar las librerías necesarias para ejecutar una serie de códigos que utilizaremos más adelante; para ello:

```
install.packages("readxl", repos = "http://cran.us.r-project.org")
install.packages("ggplot2", repos = "http://cran.us.r-project.org")
install.packages("ordinal", repos = "http://cran.us.r-project.org")
```

```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.2.3
```

```
library(ggplot2)
library(ordinal)
```

```
## Warning: package 'ordinal' was built under R version 4.2.3
```

Antes de continuar, deberemos abrir la base de datos que acompaña a este trabajo y agregarla a nuestro environment en forma de objeto; para así poder trabajar con ella. Ya que la base de datos está guardada en forma de documento de excel (formato .xlsx), correremos el siguiente código:

```
dat <- read_excel(path = "Base de datos Pileta.xlsx")
```

Lo que hemos hecho es crear un objeto llamado *dat* en el que almacenamos todos los datos proporcionados por nuestra base de datos.

Otro paso necesario antes de continuar es convertir en factores los datos proporcionados por nuestra base de datos, para lo que será necesario correr el siguiente código:

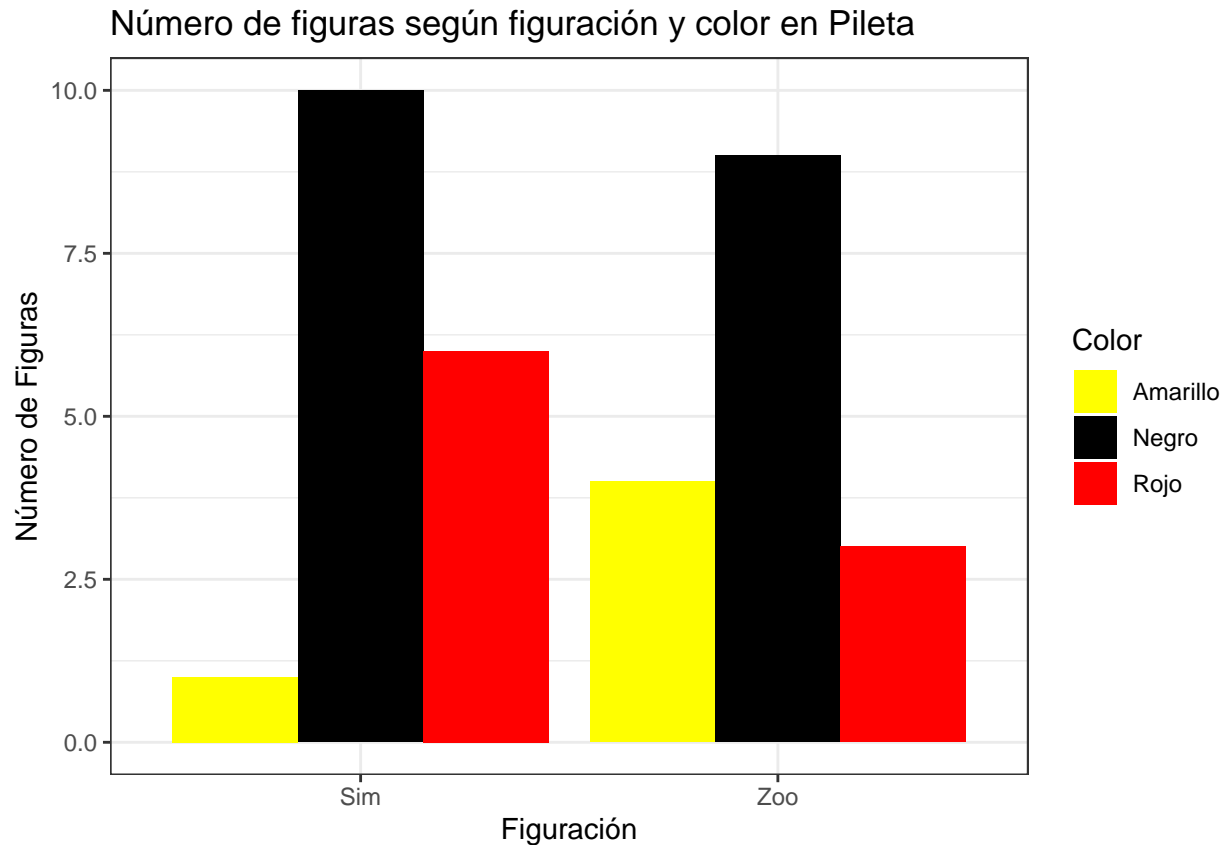
```
dat$Figuracion <- as.factor(dat$Figuracion)
dat$Color <- as.factor(dat$Color)
dat$Estilo <- as.factor(dat$Estilo)
```

Este paso es esencial, porque no podremos trabajar con nuestros datos a no ser que se encuentren en forma de factores.

3.2. Gráficas:

En una primera instancia, podemos agrupar los datos proporcionados en una gráfica, para observar a simple vista si la distribución de los datos podría proporcionarnos algún tipo de correlación estadística significativa. Para ello podemos correr el siguiente código para generar una gráfica:

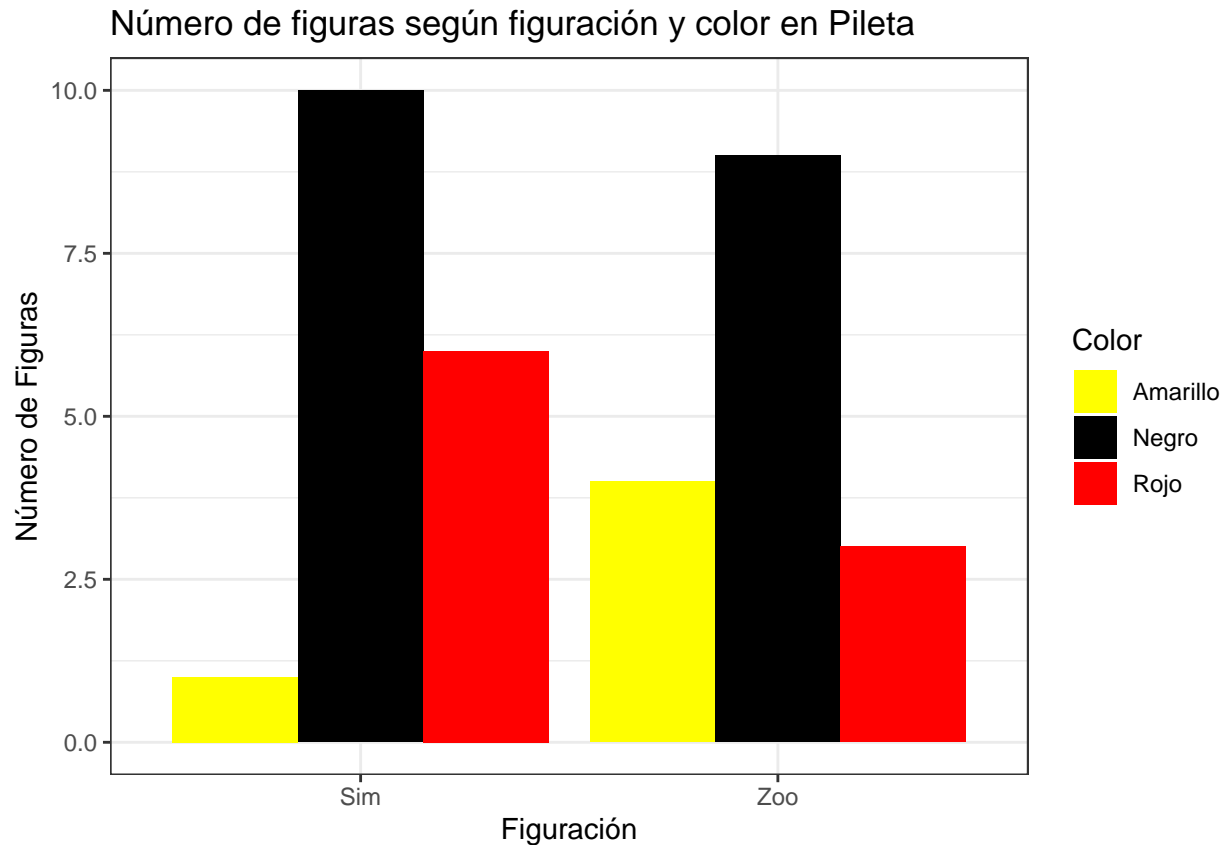
```
ggplot(data = dat, aes(x=Figuracion, fill=Color)) + geom_bar(position = "dodge") + theme_bw() +
  scale_fill_manual(values = c("Yellow", "Black", "Red")) + labs(x = "Figuración", y="Número de Figuras")
```



Como explicación al código expuesto: *ggplot* es la base del código, una función que podemos utilizar para generar gráficas. *data* es el argumento que dispone que datos se están utilizando para generar la gráfica. *aes* es el argumento que dicta que se mapeará en la gráfica, en este caso estamos diciendo que la variable *figuración* debe ir al eje X, y que será la variable *color* la que funcione como el “relleno” de las barras, es decir cómo de altas o bajas deben ser. Con el argumento *geom_bar(position = dodge)* estamos especificando cómo queremos que aparezcan las barras agrupadas, en este caso, un grupo al lado del otro. *theme_bw* está especificando la estética del gráfico, para que aparezcan las líneas más finas que vemos al fondo, por ejemplo. Con *scale_fill_manual* estamos especificando los colores de las barras del gráfico, que en este caso son: amarillo, negro y rojo. Para ello es necesario crearlas como un vector, por lo que utilizamos *c =*. Por último, con el argumento, *labs* estamos dictando las etiquetas que debe tener el gráfico; *x* dicta la etiqueta del eje X, *y* la del eje Y; y *title* el título del gráfico que aparecerá en la parte superior.

Una vez hemos comprendido esto, volvemos a nuestro gráfico:

```
ggplot(data = dat, aes(x=Figuracion, fill=Color)) + geom_bar(position = "dodge") + theme_bw() +
  scale_fill_manual(values = c("Yellow", "Black", "Red")) + labs( x= "Figuración", y="Número de Figuras")
```



Para interpretar este gráfico debemos comprender que se está representando en forma de barra la cantidad de figuras que están representadas con un color específico; además están separadas en dos grupos, a la izquierda, las tres primeras barras representan figuras encajadas en la categoría de Símbolo, mientras que las tres siguientes, a la derecha, agrupan aquellas pinturas descritas como zoomorfos. A primera vista podemos reconocer que la distribución de zoomorfos es bastante similar para los tres tipos de pintura, mientras que la de símbolos es más desigual, destacando por su cantidad las de pintura negra, y también por su escasez las de pintura amarilla.

Gracias a este primera visual de nuestros datos podríamos inferir que probablemente encontraremos una relación estadística entre la pintura de color amarillo y las representaciones de zoomorfos; y que probablemente también exista una relación entre las pinturas negras y las representaciones de símbolos. De la misma forma, también podemos inferir que será difícil encontrar una relación entre las pinturas rojas y cualquiera de las dos categorías, pues la distribución de pinturas es bastante similar entre ambos grupos.

Otro gráfico que podemos obtener, pondrá en relación los tipos de color de pintura con el estilo que presentan. Para ello utilizaremos un código muy similar al anterior, pero cambiando los datos para obtener el resultado que queremos. De esta forma, podemos correr el siguiente código:

```
ggplot(data = dat, aes(x=Estilo, fill=Color)) + geom_bar(position = "dodge") + theme_bw() +
  scale_fill_manual(values = c("Yellow", "Black", "Red")) + labs(x = "Estilo", y = "Número de Figuras", t
```



Aquí graficamos por barras los distintos tipos de estilos que encontramos en nuestra base de datos, teniendo a la izquierda el más naturalista (estilo 1), y progresando hacia la derecha hacia estilos más esquemáticos hasta llegar al estilo 5 (esquemático).

Lo que podemos observar a simple vista es una distribución en la que el amarillo ocupa los primeros puestos y desaparece prácticamente hasta un solo caso en el estilo 5; por lo que probablemente podamos encontrar una relación entre este color y el estilo naturalista. Por otro lado tanto el rojo como el negro tienen una distribución más difícil de describir, por lo que no podemos inferir a simple vista si obtendremos resultados estadísticos significativos.

3.3. Figuración vs. Color:

Lo primero que queremos comprobar en este trabajo es si existe una relación estadística significativa entre las variables figuración y color. Para ello utilizaremos un modelo conocido como *GLM* (Modelo Lineal Generalizado). Este tipo de modelo nos permite establecer relaciones entre variables cuando una de ellas no tiene una distribución normal, como es nuestro caso, en la que tiene una distribución binomial (puede ser 0 ó 1; sí o no; zoomorfo o no-zoomorfo).

Para conseguir el objetivo de contrastar la relación estadística entre estas dos variables, dividiremos nuestra base de datos por colores, para comprobar individualmente la relación existente entre uno de los tres colores con la de figuración. Para ello comprobaremos la posibilidad que tienen las pinturas de cada color de ser o no un zoomorfo.

Antes de correr el código que nos proporcionará los datos necesarios para saber si existe una correlación entre un color concreto y ser zoomorfos, tenemos que seguir una serie de pasos.

Primero creamos una tabla para nuestros datos con la función *table*; los argumentos en el interior de la función indican de dónde proceden los datos: *datFiguracion* indica que el valor X procede de la tabla de

datos *dat* de la columna *Figuracion*; y el siguiente argumento dice lo mismo pero para el eje Y y la columna *Color* de la misma base de datos.

```
table(dat$Figuracion, dat$Color)
```

```
##
##      Amarillo Negro Rojo
## Sim         1    10    6
## Zoo         4     9    3
```

Para poder trabajar con los datos proporcionados con el método explicado con anterioridad, será necesario crear nuevas columnas en nuestro dataframe, para así, poder operar con ellas. Esto lo conseguiremos gracias al siguiente código:

```
dat$Amarillo <- ifelse(dat$Color=="Amarillo", 1,0)
dat$Rojo <- ifelse(dat$Color=="Rojo", 1,0)
dat$Negro <- ifelse(dat$Color=="Negro", 1,0)
```

Lo que hemos conseguido es crear columnas individuales en base al color en nuestro dataframe, que separan los casos en función del color que presenten. Esto es posible gracias a la función `<` – que crea la nueva columna, y la función *ifelse* que transforma la columna original en valores binarios (0 ó 1) en función de si cumple la condición de ser del color indicado. Es decir, en el primer ejemplo, dará valor 1 a aquellos casos en los que el color de la pintura es amarilla. Repetimos esto para los colores Rojo y Negro. Este sería el resultado para el color amarillo:

```
dat$Amarillo
```

```
## [1] 0 0 0 1 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Por otra parte, la variable *Figuración* ya se encuentra codificada como binomial, gracias a la función *as.factor* que hemos utilizado como anterioridad; por lo que no es necesario recodificarla como hemos hecho para la variable de color.

Teniendo esto, podemos correr los modelos GLM individualmente; comparando cada color individualmente con la posibilidad o no de ser un zoomorfo. Para ello utilizaremos el siguiente código:

```
summary(glm(formula = Figuracion ~ Amarillo, data = dat, family = "binomial"))
```

```
##
## Call:
## glm(formula = Figuracion ~ Amarillo, family = "binomial", data = dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.794  -1.058  -1.058   1.302   1.302
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.2877     0.3819  -0.753   0.451
## Amarillo       1.6740     1.1815   1.417   0.157
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
## Null deviance: 45.717 on 32 degrees of freedom
## Residual deviance: 43.247 on 31 degrees of freedom
## AIC: 47.247
##
## Number of Fisher Scoring iterations: 4
```

El desglose del código utilizado es el siguiente: La función *summary* nos da un resumen de los resultados del modelo que creamos a continuación. La función *glm* es la que crea el modelo lineal generalizado que hemos explicado con antelación. El argumento *formula* es el que establece las variables respuesta y predictor, en este caso, *Figuracion* es la variable de respuesta y *Amarillo* la predictor. Es decir, estamos comprobando la posibilidad o no ser un zoomorfo (ya que zoomorfo será el factor que se utilice para estos cálculos) en base a que la pintura sea amarilla. El argumento *data* indica al modelo de dónde debe proceder la información, en este caso, procederá de nuestro dataframe *dat*. Por último el argumento *family* indica la familia de distribución de los datos del modelo; en este caso, como ha sido expuesto debe ser binomial (2 posibles factores; sí o no; 0 y 1).

De esta manera, lo que conseguimos es una serie de datos en el resumen del modelo, de los que extraeremos el p-value, que es el que nos informa sobre si hay una relación estadística significativa para estas dos variables. En este caso, dado que el p-value es menor a 0'05 obtenemos una relación estadística significativa; es decir, que sí existe una relación entre ser zoomorfo o no ser un zoomorfo y estar pintado en amarillo.

Ahora corremos los otros modelos para comprobar la misma relación pero para la pintura roja y la pintura negra; primero la roja:

```
summary(glm(formula = Figuracion ~ Rojo, data = dat, family = "binomial"))
```

```
##
## Call:
## glm(formula = Figuracion ~ Rojo, family = "binomial", data = dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2491  -1.2491  -0.9005   1.1073   1.4823
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.1671     0.4097   0.408   0.683
## Rojo          -0.8602     0.8172  -1.053   0.293
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 45.717 on 32 degrees of freedom
## Residual deviance: 44.561 on 31 degrees of freedom
## AIC: 48.561
##
## Number of Fisher Scoring iterations: 4
```

En este caso, no obtenemos una relación estadística significativa, ya que el p-value es mayor a 0.05; por lo que no podemos decir que exista una relación entre ser o no ser un zoomorfo en base a estar dibujado en color rojo.

Por último, en esta sección, las pinturas negras:


```
summary(glm(formula = Figuracion ~ Negro, data = dat, family = "binomial"))
```

```
##
## Call:
## glm(formula = Figuracion ~ Negro, family = "binomial", data = dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.177  -1.133  -1.133   1.222   1.222
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  9.671e-18  5.345e-01   0.000   1.000
## Negro       -1.054e-01  7.049e-01  -0.149   0.881
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 45.717  on 32  degrees of freedom
## Residual deviance: 45.695  on 31  degrees of freedom
## AIC: 49.695
##
## Number of Fisher Scoring iterations: 3
```

Aunque el p-value que obtenemos es menor al caso anterior, tampoco es menor a 0.05, por lo que tampoco podemos decir que exista una relación entre ser o no ser zoomorfo y estar dibujado en negro.

Dado que hemos comprobado si existe una relación entre la variable de figuración y la de color; podemos concluir que solo para el color amarillo existe una relación estadística positiva.

Comprobar si existe o no una relación entre ser o no ser símbolo en relación al color de las pinturas, debería ofrecer los mismos resultados pero algunos de los coeficientes negativos; ya que estamos repitiendo la misma acción, el mismo proceso estadístico, pero de forma inversa. Si quisieramos comprobarlo utilizaríamos este código:

```
summary(glm(formula = relevel(Figuracion, ref = "Sim") ~ Amarillo, data = dat, family = "binomial"))
```

```
##
## Call:
## glm(formula = relevel(Figuracion, ref = "Sim") ~ Amarillo, family = "binomial",
##      data = dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.794  -1.058  -1.058   1.302   1.302
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.2877    0.3819  -0.753   0.451
## Amarillo      1.6740    1.1815   1.417   0.157
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 45.717  on 32  degrees of freedom
```

```
## Residual deviance: 43.247 on 31 degrees of freedom
## AIC: 47.247
##
## Number of Fisher Scoring iterations: 4
```

Como puede comprobarse, es un código muy similar a los anteriores, pero encontramos una diferencia en el argumento *relevel* con el que le decimos que queremos que escoja el valor “Sim” de nuestra variable Figuración para hacer los cálculos; a diferencia de hasta ahora que estaba usando de manera predeterminada el valor “Zoo” de la variable. Los resultados, como han sido explicados, son idénticos al modelo anterior, solo que algunos de los coeficientes son negativos. El p-value sigue siendo menor a 0.05 por lo que existe una relación estadística significativa entre ser o no ser símbolo y estar dibujado en color amarillo.

3.4. Odds ratio

Una vez sabemos que hay una relación estadística significativa entre el color amarillo y la variable figuración, la siguiente pregunta que nos surge, es de qué tipo es esta relación; es decir cómo están relacionadas estadísticamente.

Para conseguir este objetivo podemos usar los *odds.ratio* (razón de posibilidades) que es un tipo de medida estadística empleada en modelos de regresión lineal (como el que estamos utilizando). Lo que estaremos calculando serán las posibilidades de éxito que existen entre la variable de respuesta y la predictora.

Para conseguir este valor en nuestro caso, haremos lo siguiente; primero convertir nuestro modelo en un objeto del que después podamos extraer datos.

```
modeloA <- glm(formula = Figuracion ~ Amarillo, data = dat, family = "binomial")
```

Ahora utilizaremos la función *coefficients* sobre nuestro modelo; lo cual extraerá los coeficientes de dicho modelo; pero para obtener los *odds.ratio* deberemos incluir el argumento *exp*. Lo que hace este último argumento es exponenciar los valores de los coeficientes y con ello obtener los valores de *odds.ratio* del modelo. En nuestro caso correremos el siguiente código:

```
exp(coefficients(modeloA))
```

```
## (Intercept)      Amarillo
##      0.750000      5.333333
```

El resultado obtenido se interpreta así: la probabilidad de ser un zoomorfo es 7 veces mayor que la de ser un símbolo para las figuras de pigmentación amarilla. La posibilidad es esta porque el valor es positivo, y el factor “Zoo” es el que es positivo en nuestro modelo. Si el resultado fuese negativo (-7.27) indicaría que la probabilidad de ser símbolo es 7 veces mayor que la de ser un zoomorfo para las pinturas amarillas.

Estos mismos pasos podrían seguirse para los otros modelos, pero ya que no existe una relación estadística significativa para los colores rojo y negro en relación a la figuración, tampoco obtendremos un *odds.ratio* que indique una relación estadística similar a la del color amarillo. Si quisieramos hacerlo, por ejemplo para el color rojo, tan solo tendríamos que correr el siguiente código:

```
modeloR <- glm(formula = Figuracion ~ Rojo, data = dat, family = "binomial")
exp(coefficients(modeloR))
```

```
## (Intercept)      Rojo
##      1.181818      0.4230769
```

Lo que nos están indicando estos resultados es que la probabilidad de ser un zoomorfo es 0'8 veces mayor que la de ser un símbolo para el color rojo; lo cuál no es demasiado significativo y por tanto no tiene importancia estadística.

Lo mismo ocurre para el color negro:

```
modeloN <- glm(formula = Figuracion ~ Negro, data = dat, family = "binomial")
exp(coefficients(modeloN))
```

```
## (Intercept)      Negro
##          1.0         0.9
```

Aquí podemos interpretar que la probabilidad de ser un zoomorfo es 0'42 veces mayor que la de ser un símbolo para el color negro; lo cuál tampoco tiene importancia estadística.

3.5. Estilo vs. Color

Otra cuestión que podemos plantear dados los datos propuestos es si existe una relación entre el color y el estilo. Ya que la variable de estilo es una de carácter ordinal, no podremos utilizar los mismos modelos que hemos usado hasta ahora. Para esta tarea utilizaremos modelos CLM (Cumulative Link Models), que son modelos preparados para manejar datos ordinales. Ya que también es un tipo de modelo de regresión binaria, el tipo de datos que nos proporciona así como su interpretación es bastante similar a lo que hemos visto hasta ahora.

Dicho esto, correremos un código muy similar a los anteriores, pero usando en este caso la función *clm*. Así para el color amarillo tenemos:

```
summary(clm(formula = Estilo ~ Amarillo, data=dat))
```

```
## formula: Estilo ~ Amarillo
## data:      dat
##
## link threshold nobs logLik AIC   niter max.grad cond.H
## logit flexible 33   -37.17 84.33 7(2)  2.12e-09 3.0e+02
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## Amarillo   -3.279      1.280   -2.562   0.0104 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
##             Estimate Std. Error z value
## 1|2   -3.9859      1.1095   -3.593
## 2|3   -1.4275      0.4618   -3.091
## 3|4   -0.4687      0.3808   -1.231
## 4|5   -0.3358      0.3773   -0.890
```

Como en los otros casos, ya que el p-value que se nos ofrece es menor a 0.05, tenemos una relación estadística significativa; lo cual nos dice que sí que hay una relación entre que el estilo y el color amarillo.

Para el color rojo y para el negro no obtenemos resultados significativos; el p-value de sus resultados es mayor a 0.05, así que no tenemos una relación estadística significativa, que implica que no existe una relación entre el estilo y los colores rojo y negro. El código que utilizamos para obtener estos datos es el siguiente:

```
summary(clm(formula = Estilo ~ Rojo, data=dat))
```

```
## formula: Estilo ~ Rojo
## data:      dat
##
## link threshold nobs logLik AIC   niter max.grad cond.H
## logit flexible  33   -40.67 91.34 6(2)  3.32e-08 1.2e+02
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## Rojo    0.9551     0.7854   1.216   0.224
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 1|2 -2.54408     0.74429  -3.418
## 2|3 -0.74518     0.43123  -1.728
## 3|4  0.07679     0.40546   0.189
## 4|5  0.20159     0.40593   0.497
```

```
summary(clm(formula = Estilo ~ Negro, data=dat))
```

```
## formula: Estilo ~ Negro
## data:      dat
##
## link threshold nobs logLik AIC   niter max.grad cond.H
## logit flexible  33   -41.22 92.44 6(2)  3.27e-08 2.1e+02
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## Negro    0.4679     0.6733   0.695   0.487
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 1|2 -2.4807     0.8136  -3.049
## 2|3 -0.6977     0.5594  -1.247
## 3|4  0.1073     0.5410   0.198
## 4|5  0.2288     0.5402   0.424
```

De igual manera que lo hemos hecho en el apartado anterior, podemos calcular los Odds Ratio de estos modelos para que nos informen sobre el tipo de relación que mantienen las variables. Ya que no existe relación estadística para los colores rojo y negro, solo lo calcularemos para el color Amarillo. El código que utilizamos es el siguiente:

```
modeloA2 <- summary(clm(formula = Estilo ~ Amarillo, data=dat))
exp(coefficients(modeloA2))
```

```
##      Estimate Std. Error    z value Pr(>|z|)
## 1|2    0.01857515  3.032740 0.02752514 1.000327
## 2|3    0.23991430  1.586957 0.04545724 1.001997
## 3|4    0.62579899  1.463460 0.29203354 1.244041
## 4|5    0.71474235  1.458346 0.41061919 1.452692
## Amarillo 0.03765499  3.597158 0.07717676 1.010472
```

Como podemos comprobar en los resultados, ningún pvalue es menor a 0'05, por lo que ninguno de nuestros resultados es estadísticamente significativo, y no podemos comprobar específicamente cómo está relacionado el color Amarillo a la variable de Estilo. Que obtengamos resultados no significativos podría deberse a que los datos están demasiado correlacionados o a que el tamaño de la muestra es pequeño. La segunda opción es más probable porque nuestra muestra solo cuenta con 33 casos.

4. Conclusiones

Como ha sido expuesto, las relaciones estadísticas descubiertas para las pinturas de la cueva de La Pileta, son ciertamente escasas. Con respecto al color de las Pinturas solo podemos saber que existe una relación estadística significativa entre el color amarillo y los zoomorfos; habiendo una probabilidad 7 veces mayor de que una figura amarilla sea un zoomorfo a que no lo sea. También sabemos que existe una relación entre el color amarillo y la variable de Estilo; pero no podemos saber cómo se relacionan, pues no obtenemos relaciones significativa en sus odds ratio.

El resto de los resultados estadísticos expuestos no son significativos; es decir, no encontramos una relación entre el color rojo y negro en las pinturas con la figura y el estilo que están presentando. Dados estos resultados, podemos corroborar que esto es lo que encontramos en la cueva gracias a los trabajos de algunos investigadores como (Breuil, 1914; Dams, 1978 y Sanchidrián y Muñoz 1990). Las pinturas amarillas eran las más probables a tener una correlación estadística con alguna de las otras dos variables, ya que, en esencia, es un grupo conformado mayoritariamente por figuras zoomorfas bastante realistas. Respecto a los otros dos colores, ambos están muy utilizados para representar tanto zoomorfos como símbolos, por lo que era difícil que se pudiera obtener una relación estadística entre alguno de los factores propuestos. Pero también debemos tener en cuenta el tamaño de la muestra, que, por restricciones de tiempo, debe ser acotada a unos 30-40 casos, de los más de 1200 que encontramos en Pileta; quizá de contar con una base de datos más completa podríamos haber conseguido resultados más significativos.

Con lo expuesto, podemos concluir que, aunque tengamos una muestra muy pequeña, con ella ya podemos establecer algunas relaciones estadísticas en relación al color amarillo y las variables de Estilo y Figuración; y aunque no podamos obtener resultados significativos para los otros dos colores, esto no es menos relevante, pues es la realidad que encontramos en la cueva es una cantidad similar de casos en los que los colores rojo y negro se utilizan para dibujar zoomorfos y símbolos de diferentes estilos.

5. Bibliografía:

Cortés Sánchez, M. y Simón Vallejo, M. D. (2007). “La Pileta (Benaolán, Málaga) cien años después. Aportaciones al conocimiento de su secuencia arqueológica.” SAGVNTVM: Papeles del Laboratorio de Arqueología de Valencia, 39, p. 45-64.

Sanchidrián Torti, J.L. (1997). Propuesta de la secuencia figurativa en la Cueva de la Pileta. Centre d'investigacions arqueològiques sèrie monogràfica, p. 411-430.

Sanchidrián Torti, J. L. (2001). Manual de arte prehistórico. Grupo Planeta (GBS).

- Consultas web para código en R:

ggplot: https://rpubs.com/hllinas/R_Barras_ggplot_univariada

GLM: <https://rpubs.com/JessicaP/459130>

Odds.Ratio: <https://es.stackoverflow.com/questions/347140/la-funci%C3%B3n-glm-genera-y-almacena-los-odd-ratio-del-modelo-estimado>

CLM: <https://community.rstudio.com/t/cumulative-link-model-in-r/67422> <https://rdr.io/cran/ordinal/man/clm.html>