EL GWAS :D

Santiago Orozco Barrera & Gabriel Ramírez Vilchis

GWAS

Desarrollo

```
# Crear data frames con los datos de los SNPs
rs4988235_observed_counts <- data.frame(brown_or_other = c(21, 55, 76),
                        blue_or_green = c(10, 12, 22), total = c(31, 67, 98),
                        row.names = c("A", "G", "total"))
rs4988235_observed_counts
```

```
      brown_or_other blue_or_green total
A                 21            10    31
G                 55            12    67
total             76            22    98
```

```
rs7495174_observed_counts <- data.frame(brown_or_other = c(55, 21, 76),
                        blue_or_green = c(22, 0, 22), total = c(77, 21, 98),
                        row.names = c("A", "G", "total"))
rs7495174_observed_counts
```

```
      brown_or_other blue_or_green total
A                 55            22    77
G                 21             0    21
total             76            22    98
```

```
rs713598_observed_counts <- data.frame(brown_or_other = c(32, 44, 76),
                        blue_or_green = c(16, 6, 22), total = c(48, 50, 98),
                        row.names = c("C", "G", "total"))
rs713598_observed_counts
```

```
      brown_or_other blue_or_green total
C                 32            16    48
G                 44             6    50
total             76            22    98
```

```
rs17822931_observed_counts <- data.frame(brown_or_other = c(41, 35, 76),
                         blue_or_green = c(16, 6, 22), total = c(57, 41, 98),
                         row.names = c("C", "T", "total"))
rs17822931_observed_counts
```

```
      brown_or_other blue_or_green total
C                 41            16    57
T                 35             6    41
total             76            22    98
```

```
rs4481887_observed_counts <- data.frame(brown_or_other = c(21, 55, 76),
                         blue_or_green = c(4, 18, 22), total = c(25, 73, 98),
                         row.names = c("A", "G", "total"))
rs4481887_observed_counts
```

```
      brown_or_other blue_or_green total
A                 21             4    25
G                 55            18    73
total             76            22    98
```

```
# Calcular frecuencias observadas
rs4988235_observed_frequencies <-
  rs4988235_observed_counts / rs4988235_observed_counts["total", "total"]
rs4988235_observed_frequencies
```

```
      brown_or_other blue_or_green     total
A          0.2142857     0.1020408 0.3163265
G          0.5612245     0.1224490 0.6836735
total      0.7755102     0.2244898 1.0000000
```

```
rs7495174_observed_frequencies <-
  rs7495174_observed_counts / rs7495174_observed_counts["total", "total"]
rs7495174_observed_frequencies
```

```
      brown_or_other blue_or_green     total
A          0.5612245     0.2244898 0.7857143
G          0.2142857     0.0000000 0.2142857
total      0.7755102     0.2244898 1.0000000
```

```
rs713598_observed_frequencies <-
  rs713598_observed_counts / rs713598_observed_counts["total", "total"]
rs713598_observed_frequencies
```

```
      brown_or_other blue_or_green     total
C          0.3265306    0.16326531 0.4897959
G          0.4489796    0.06122449 0.5102041
total      0.7755102    0.22448980 1.0000000
```

```
rs17822931_observed_frequencies <-
  rs17822931_observed_counts / rs17822931_observed_counts["total", "total"]
rs17822931_observed_frequencies
```

```
      brown_or_other blue_or_green     total
C          0.4183673    0.16326531 0.5816327
T          0.3571429    0.06122449 0.4183673
total      0.7755102    0.22448980 1.0000000
```

```
rs4481887_observed_frequencies <-
  rs4481887_observed_counts / rs4481887_observed_counts["total", "total"]
rs4481887_observed_frequencies
```

```
      brown_or_other blue_or_green    total
A          0.2142857    0.04081633 0.255102
G          0.5612245    0.18367347 0.744898
total      0.7755102    0.22448980 1.000000
```

```
# rs4988235
## Calcular frecuencias esperadas
rs4988235_expected_frequencies <- data.frame(
  brown_or_other = c(rs4988235_observed_frequencies["total", "brown_or_other"] *
                       rs4988235_observed_frequencies["A", "total"],
                     rs4988235_observed_frequencies["total", "brown_or_other"] *
                       rs4988235_observed_frequencies["G", "total"]),
  blue_or_green = c(rs4988235_observed_frequencies["total", "blue_or_green"] *
```

```
                      rs4988235_observed_frequencies["A", "total"],
                    rs4988235_observed_frequencies["total", "blue_or_green"] *
                      rs4988235_observed_frequencies["G", "total"]),
  row.names = c("A", "G"))

## Agregar la fila de totales
rs4988235_expected_frequencies <- rbind(rs4988235_expected_frequencies,
  total = rs4988235_expected_frequencies["A", ] +
    rs4988235_expected_frequencies["G", ])

## Agregar la columna de totales
rs4988235_expected_frequencies <- cbind(rs4988235_expected_frequencies,
  total = rs4988235_expected_frequencies[, "brown_or_other"] +
    rs4988235_expected_frequencies[, "blue_or_green"])

## Visualizar frecuencias esperadas
rs4988235_expected_frequencies
```

```
      brown_or_other blue_or_green     total
A          0.2453145    0.07101208 0.3163265
G          0.5301958    0.15347772 0.6836735
total      0.7755102    0.22448980 1.0000000
```

```
# rs7495174
## Calcular frecuencias esperadas
rs7495174_expected_frequencies <- data.frame(
  brown_or_other = c(rs7495174_observed_frequencies["total", "brown_or_other"] *
                       rs7495174_observed_frequencies["A", "total"],
                     rs7495174_observed_frequencies["total", "brown_or_other"] *
                       rs7495174_observed_frequencies["G", "total"]),
  blue_or_green = c(rs7495174_observed_frequencies["total", "blue_or_green"] *
                      rs7495174_observed_frequencies["A", "total"],
                    rs7495174_observed_frequencies["total", "blue_or_green"] *
                      rs7495174_observed_frequencies["G", "total"]),
  row.names = c("A", "G"))

## Agregar la fila de totales
rs7495174_expected_frequencies <- rbind(rs7495174_expected_frequencies,
  total = rs7495174_expected_frequencies["A", ] +
    rs7495174_expected_frequencies["G", ])
```

```r
## Agregar la columma de totales
rs7495174_expected_frequencies <- cbind(rs7495174_expected_frequencies,
  total = rs7495174_expected_frequencies[, "brown_or_other"] +
    rs7495174_expected_frequencies[, "blue_or_green"])

## Visualizar frecuencias esperadas
rs7495174_expected_frequencies
```

```
      brown_or_other blue_or_green     total
A          0.6093294    0.17638484 0.7857143
G          0.1661808    0.04810496 0.2142857
total      0.7755102    0.22448980 1.0000000
```

```r
# rs713598
## Calcular frecuencias esperadas
rs713598_expected_frequencies <- data.frame(
  brown_or_other = c(rs713598_observed_frequencies["total", "brown_or_other"] *
                       rs713598_observed_frequencies["C", "total"],
                     rs713598_observed_frequencies["total", "brown_or_other"] *
                       rs713598_observed_frequencies["G", "total"]),
  blue_or_green = c(rs713598_observed_frequencies["total", "blue_or_green"] *
                      rs713598_observed_frequencies["C", "total"],
                    rs713598_observed_frequencies["total", "blue_or_green"] *
                      rs713598_observed_frequencies["G", "total"]),
  row.names = c("C", "G"))

## Agregar la fila de totales
rs713598_expected_frequencies <- rbind(rs713598_expected_frequencies,
  total = rs713598_expected_frequencies["C", ] +
    rs713598_expected_frequencies["G", ])

## Agregar la columma de totales
rs713598_expected_frequencies <- cbind(rs713598_expected_frequencies,
  total = rs713598_expected_frequencies[, "brown_or_other"] +
    rs713598_expected_frequencies[, "blue_or_green"])

## Visualizar frecuencias esperadas
rs713598_expected_frequencies
```

```
      brown_or_other blue_or_green     total
C          0.3798417     0.1099542 0.4897959
```

```
G           0.3956685     0.1145356 0.5102041
total       0.7755102     0.2244898 1.0000000
```

```r
# rs17822931
## Calcular frecuencias esperadas
rs17822931_expected_frequencies <- data.frame(
  brown_or_other = c(rs17822931_observed_frequencies["total", "brown_or_other"] *
                       rs17822931_observed_frequencies["C", "total"],
                     rs17822931_observed_frequencies["total", "brown_or_other"] *
                       rs17822931_observed_frequencies["T", "total"]),
  blue_or_green = c(rs17822931_observed_frequencies["total", "blue_or_green"] *
                      rs17822931_observed_frequencies["C", "total"],
                    rs17822931_observed_frequencies["total", "blue_or_green"] *
                      rs17822931_observed_frequencies["T", "total"]),
  row.names = c("C", "T"))

## Agregar la fila de totales
rs17822931_expected_frequencies <- rbind(rs17822931_expected_frequencies,
  total = rs17822931_expected_frequencies["C", ] +
    rs17822931_expected_frequencies["T", ])

## Agregar la columna de totales
rs17822931_expected_frequencies <- cbind(rs17822931_expected_frequencies,
  total = rs17822931_expected_frequencies[, "brown_or_other"] +
    rs17822931_expected_frequencies[, "blue_or_green"])

## Visualizar frecuencias esperadas
rs17822931_expected_frequencies
```

```
      brown_or_other blue_or_green     total
C          0.4510621     0.1305706 0.5816327
T          0.3244481     0.0939192 0.4183673
total      0.7755102     0.2244898 1.0000000
```

```r
# rs4481887
## Calcular frecuencias esperadas
rs4481887_expected_frequencies <- data.frame(
  brown_or_other = c(rs4481887_observed_frequencies["total", "brown_or_other"] *
                       rs4481887_observed_frequencies["A", "total"],
                     rs4481887_observed_frequencies["total", "brown_or_other"] *
                       rs4481887_observed_frequencies["G", "total"]),
  blue_or_green = c(rs4481887_observed_frequencies["total", "blue_or_green"] *
```

```
                    rs4481887_observed_frequencies["A", "total"],
                  rs4481887_observed_frequencies["total", "blue_or_green"] *
                    rs4481887_observed_frequencies["G", "total"]),
  row.names = c("A", "G"))

## Agregar la fila de totales
rs4481887_expected_frequencies <- rbind(rs4481887_expected_frequencies,
  total = rs4481887_expected_frequencies["A", ] +
    rs4481887_expected_frequencies["G", ])

## Agregar la columna de totales
rs4481887_expected_frequencies <- cbind(rs4481887_expected_frequencies,
  total = rs4481887_expected_frequencies[, "brown_or_other"] +
    rs4481887_expected_frequencies[, "blue_or_green"])

## Visualizar frecuencias esperadas
rs4481887_expected_frequencies
```

```
      brown_or_other blue_or_green    total
A          0.1978342    0.05726781 0.255102
G          0.5776760    0.16722199 0.744898
total      0.7755102    0.22448980 1.000000
```

```
# Calcular conteos esperados
rs4988235_expected_counts <- rs4988235_expected_frequencies *
  rs4988235_observed_counts["total", "total"]
rs4988235_expected_counts
```

```
      brown_or_other blue_or_green total
A           24.04082      6.959184    31
G           51.95918     15.040816    67
total       76.00000     22.000000    98
```

```
rs7495174_expected_counts <- rs7495174_expected_frequencies *
  rs7495174_observed_counts["total", "total"]
rs7495174_expected_counts
```

```
      brown_or_other blue_or_green total
A           59.71429     17.285714    77
G           16.28571      4.714286    21
total       76.00000     22.000000    98
```

```
rs713598_expected_counts <- rs713598_expected_frequencies *
  rs713598_observed_counts["total", "total"]
rs713598_expected_counts
```

```
      brown_or_other blue_or_green total
C           37.22449      10.77551    48
G           38.77551      11.22449    50
total       76.00000      22.00000    98
```

```
rs17822931_expected_counts <- rs17822931_expected_frequencies *
  rs17822931_observed_counts["total", "total"]
rs17822931_expected_counts
```

```
      brown_or_other blue_or_green total
C           44.20408     12.795918    57
T           31.79592      9.204082    41
total       76.00000     22.000000    98
```

```
rs4481887_expected_counts <- rs4481887_expected_frequencies *
  rs4481887_observed_counts["total", "total"]
rs4481887_expected_counts
```

```
      brown_or_other blue_or_green total
A           19.38776      5.612245    25
G           56.61224     16.387755    73
total       76.00000     22.000000    98
```

```
# Retirar totales de los conteos observados
rs4988235_observed_counts_reduced <- rs4988235_observed_counts[-3, -3]
rs7495174_observed_counts_reduced <- rs7495174_observed_counts[-3, -3]
rs713598_observed_counts_reduced <- rs713598_observed_counts[-3, -3]
rs17822931_observed_counts_reduced <- rs17822931_observed_counts[-3, -3]
rs4481887_observed_counts_reduced <- rs4481887_observed_counts[-3, -3]

# Aplicar el test de Fisher
fisher.test(rs4988235_observed_counts_reduced)
```

```
    Fisher's Exact Test for Count Data
```

```
data:  rs4988235_observed_counts_reduced
p-value = 0.126
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.1551857 1.3878965
sample estimates:
odds ratio
 0.4621524
```

```
fisher.test(rs7495174_observed_counts_reduced)
```

```
    Fisher's Exact Test for Count Data

data:  rs7495174_observed_counts_reduced
p-value = 0.002897
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.000000 0.540205
sample estimates:
odds ratio
         0
```

```
fisher.test(rs713598_observed_counts_reduced)
```

```
    Fisher's Exact Test for Count Data

data:  rs713598_observed_counts_reduced
p-value = 0.01522
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.07939309 0.84575962
sample estimates:
odds ratio
 0.2763657
```

```
fisher.test(rs17822931_observed_counts_reduced)
```

```
    Fisher's Exact Test for Count Data

data:  rs17822931_observed_counts_reduced
p-value = 0.1444
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.1275718 1.3549809
sample estimates:
odds ratio
 0.4428437
```

```
fisher.test(rs4481887_observed_counts_reduced)
```

```
    Fisher's Exact Test for Count Data

data:  rs4481887_observed_counts_reduced
p-value = 0.4218
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.4822422 7.7563361
sample estimates:
odds ratio
  1.709419
```

```
# ODDS RATIOS
## rs4988235
### Brown or other
#### Calcular probabilidades
brown_A_prob <- rs4988235_observed_counts["A", "brown_or_other"] /
  rs4988235_observed_counts["A", "total"]
brown_G_prob <- rs4988235_observed_counts["G", "brown_or_other"] /
  rs4988235_observed_counts["G", "total"]
#### Calcular odds
odds_A <- brown_A_prob / (1 - brown_A_prob)
odds_G <- brown_G_prob / (1 - brown_G_prob)
#### Calcular odds ratio
odds_ratio_A_G <- odds_A / odds_G
odds_ratio_A_G
```

```
[1] 0.4581818
```

```r
### Blue or green
#### Calcular probabilidades
blue_A_prob <- rs4988235_observed_counts["A", "blue_or_green"] /
  rs4988235_observed_counts["A", "total"]
blue_G_prob <- rs4988235_observed_counts["G", "blue_or_green"] /
  rs4988235_observed_counts["G", "total"]
#### Calcular odds
odds_A <- blue_A_prob / (1 - blue_A_prob)
odds_G <- blue_G_prob / (1 - blue_G_prob)
#### Calcular odds ratio
odds_ratio_A_G <- odds_A / odds_G
odds_ratio_A_G
```

```
[1] 2.18254
```

```r
## rs7495174
### Brown or other
#### Calcular probabilidades
brown_A_prob <- rs7495174_observed_counts["A", "brown_or_other"] /
  rs7495174_observed_counts["A", "total"]
brown_G_prob <- rs7495174_observed_counts["G", "brown_or_other"] /
  rs7495174_observed_counts["G", "total"]
#### Calcular odds
odds_A <- brown_A_prob / (1 - brown_A_prob)
odds_G <- brown_G_prob / (1 - brown_G_prob)
#### Calcular odds ratio
odds_ratio_A_G <- odds_A / odds_G
odds_ratio_A_G
```

```
[1] 0
```

```r
### Blue or green
#### Calcular probabilidades
blue_A_prob <- rs7495174_observed_counts["A", "blue_or_green"] /
  rs7495174_observed_counts["A", "total"]
blue_G_prob <- rs7495174_observed_counts["G", "blue_or_green"] /
  rs7495174_observed_counts["G", "total"]
#### Calcular odds
odds_A <- blue_A_prob / (1 - blue_A_prob)
odds_G <- blue_G_prob / (1 - blue_G_prob)
#### Calcular odds ratio
```

```r
odds_ratio_A_G <- odds_A / odds_G
odds_ratio_A_G
```

```
[1] Inf
```

```r
## rs713598
### Brown or other
#### Calcular probabilidades
brown_C_prob <- rs713598_observed_counts["C", "brown_or_other"] /
  rs713598_observed_counts["C", "total"]
brown_G_prob <- rs713598_observed_counts["G", "brown_or_other"] /
  rs713598_observed_counts["G", "total"]
#### Calcular odds
odds_C <- brown_C_prob / (1 - brown_C_prob)
odds_G <- brown_G_prob / (1 - brown_G_prob)
#### Calcular odds ratio
odds_ratio_C_G <- odds_C / odds_G
odds_ratio_C_G
```

```
[1] 0.2727273
```

```r
### Blue or green
#### Calcular probabilidades
blue_C_prob <- rs713598_observed_counts["C", "blue_or_green"] /
  rs713598_observed_counts["C", "total"]
blue_G_prob <- rs713598_observed_counts["G", "blue_or_green"] /
  rs713598_observed_counts["G", "total"]
#### Calcular odds
odds_C <- blue_C_prob / (1 - blue_C_prob)
odds_G <- blue_G_prob / (1 - blue_G_prob)
#### Calcular odds ratio
odds_ratio_C_G <- odds_C / odds_G
odds_ratio_C_G
```

```
[1] 3.666667
```

```r
## rs17822931
### Brown or other
#### Calcular probabilidades
brown_C_prob <- rs17822931_observed_counts["C", "brown_or_other"] /
```

```
  rs17822931_observed_counts["C", "total"]
brown_T_prob <- rs17822931_observed_counts["T", "brown_or_other"] /
  rs17822931_observed_counts["T", "total"]
#### Calcular odds
odds_C <- brown_C_prob / (1 - brown_C_prob)
odds_T <- brown_T_prob / (1 - brown_T_prob)
#### Calcular odds ratio
odds_ratio_C_T <- odds_C / odds_T
odds_ratio_C_T
```

```
[1] 0.4392857
```

```
### Blue or green
#### Calcular probabilidades
blue_C_prob <- rs17822931_observed_counts["C", "blue_or_green"] /
  rs17822931_observed_counts["C", "total"]
blue_T_prob <- rs17822931_observed_counts["T", "blue_or_green"] /
  rs17822931_observed_counts["T", "total"]
#### Calcular odds
odds_C <- blue_C_prob / (1 - blue_C_prob)
odds_T <- blue_T_prob / (1 - blue_T_prob)
#### Calcular odds ratio
odds_ratio_C_T <- odds_C / odds_T
odds_ratio_C_T
```

```
[1] 2.276423
```

```
## rs4481887
### Brown or other
#### Calcular probabilidades
brown_A_prob <- rs4481887_observed_counts["A", "brown_or_other"] /
  rs4481887_observed_counts["A", "total"]
brown_G_prob <- rs4481887_observed_counts["G", "brown_or_other"] /
  rs4481887_observed_counts["G", "total"]
#### Calcular odds
odds_A <- brown_A_prob / (1 - brown_A_prob)
odds_G <- brown_G_prob / (1 - brown_G_prob)
#### Calcular odds ratio
odds_ratio_A_G <- odds_A / odds_G
odds_ratio_A_G
```

```
[1] 1.718182
```

```r
### Blue or green
#### Calcular probabilidades
blue_A_prob <- rs4481887_observed_counts["A", "blue_or_green"] /
  rs4481887_observed_counts["A", "total"]
blue_G_prob <- rs4481887_observed_counts["G", "blue_or_green"] /
  rs4481887_observed_counts["G", "total"]
#### Calcular odds
odds_A <- blue_A_prob / (1 - blue_A_prob)
odds_G <- blue_G_prob / (1 - blue_G_prob)
#### Calcular odds ratio
odds_ratio_A_G <- odds_A / odds_G
odds_ratio_A_G
```

```
[1] 0.5820106
```