

Trabajo 1 datos categóricos

1. Prueba exacta de Fisher:

Se introduce la prueba exacta de Fisher como una prueba que ayuda a probar independencia entre variables a partir de tablas de contingencia, para esta ilustración se realizará un ejemplo mediante datos simulados asociados a la tasa de empleados en los egresados de dos universidades ficticias:

Primeramente se generan dos muestras de 50 estudiantes de dos universidades, considere las siguientes variables aleatorias:

X_i : Estado de empleado del i-esimo egresado de la universidad X.

Y_i : Estado de empleado del i-esimo egresado de la universidad Y.

Se tiene que $\sum X_i \sim \text{Binom}(50, 0.6)$ y $\sum Y_i \sim \text{Binom}(50, 0.9)$.

Se obtiene así los siguientes resultados simulados:

	Universidad X	Universidad Y
Desempleado	20	1
Empleado	30	49

Bajo el esquema de la prueba de Fisher se plantea el siguiente juego de hipótesis:

H_0 : No existe asociación entre la pertenencia a una determinada universidad y la empleabilidad de sus egresados.

H_a : Existe evidencia de asociación entre la pertenencia a una determinada universidad y la empleabilidad de sus egresados.

Se plantea la prueba exacta de Fisher mediante el la función `fisher.test()` del paquete stats de R:

Fisher's Exact Test for Count Data

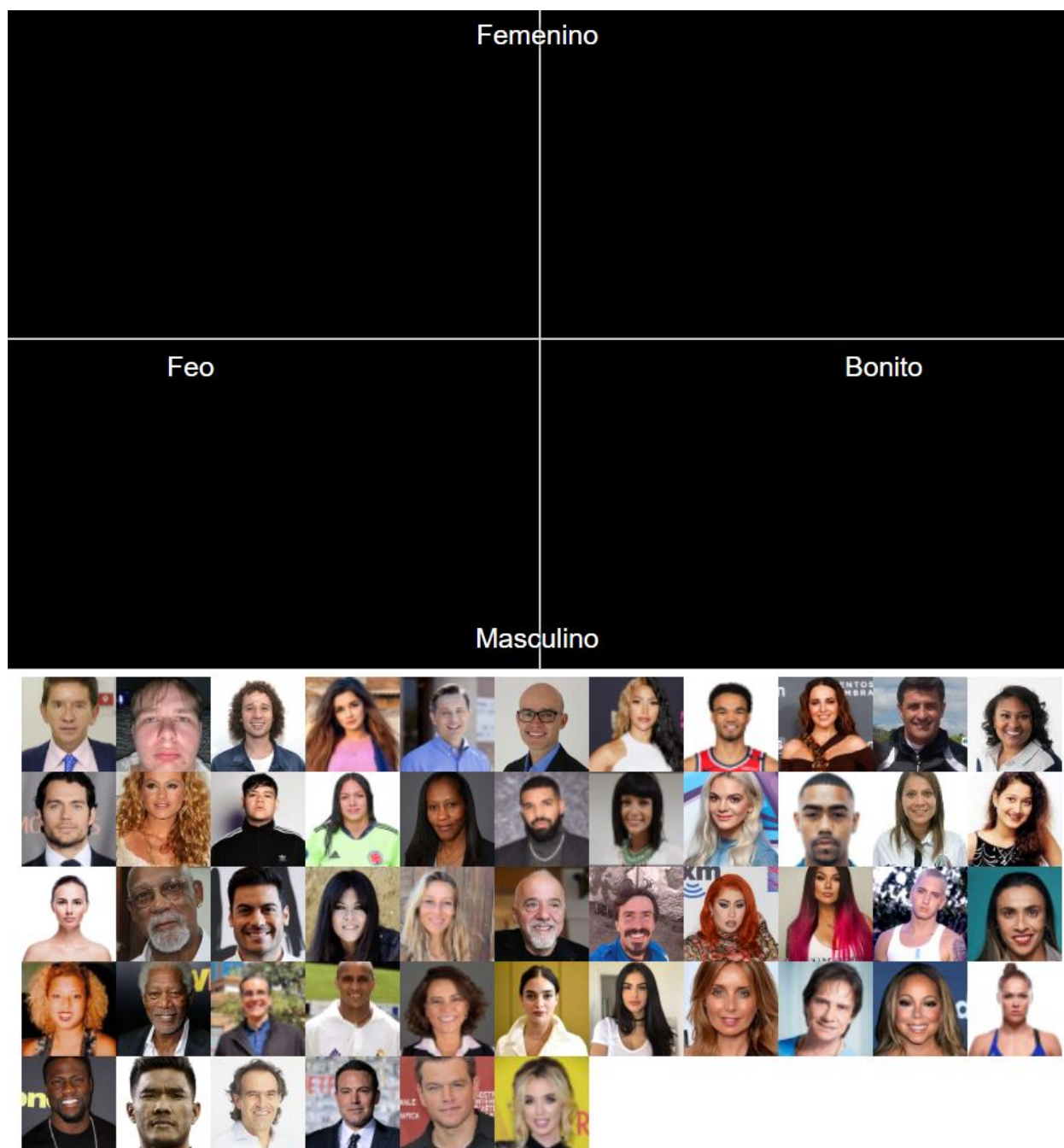
```
data:  datos_simulados
p-value = 0.000002374
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 4.612654 1370.909354
sample estimates:
odds ratio
31.73501
```

Ya que $V_p < 0.05$ se concluye con una significancia $\alpha = 0.05$ que según la prueba exacta de fisher se tiene evidencia de que se tiene asociación entre pertenecer a una universidad u otra y ser empleado o desempleado estando recién egresado, resultado consecuente con la forma en que se planteó la simulación de los datos.

Es importante resaltar en este caso que los datos son simulados mediante supuestos distribucionales binomiales y que la prueba exacta de fisher usa un esquema distribucional hipergeométrico. Aún así los resultados obtenidos en la prueba son consistentes con lo que se planteó inicialmente.

4. Lectura de datos de imagenes

Se obtienen 50 imagenes en donde todas las personas están de frente, usando luz medianamente natural y sin portar ningún accesorio, 25 de personas con rasgos femeninos y 25 de personas con rasgos masculinos. Se pregunta a 10 personas (5 hombres y 5 mujeres) acerca de a cuántas personas considera como bonitas y a cuantas personas considera como feas, para ello se utiliza la herramienta interactiva 'tiermaker' en dónde los usuarios pueden crear tablas de clasificación de fotografías:



Obteniendo así los siguientes resultados por genero plasmados en tablas 2x2:

	Bonito	Feo
<i>Mujer</i>	15	10
<i>Hombre</i>	7	18

	Bonito	Feo
<i>Mujer</i>	15	10
<i>Hombre</i>	11	14

	Bonito	Feo
<i>Mujer</i>	17	8
<i>Hombre</i>	9	16

	Bonito	Feo
<i>Mujer</i>	9	16
<i>Hombre</i>	13	12

	Bonito	Feo
<i>Mujer</i>	5	20
<i>Hombre</i>	13	12

	Bonito	Feo
<i>Mujer</i>	13	12
<i>Hombre</i>	12	13

	Bonito	Feo
<i>Mujer</i>	11	14
<i>Hombre</i>	11	14

	Bonito	Feo
<i>Mujer</i>	4	21
<i>Hombre</i>	6	19

	Bonito	Feo
<i>Mujer</i>	3	22
<i>Hombre</i>	16	9

	Bonito	Feo
<i>Mujer</i>	6	19
<i>Hombre</i>	7	18

De los resultados obtenidos se generan tres tablas 2x2, una general con el fin de obtener rasgos generales, una para las clasificaciones realizadas por las mujeres y otra para las clasificaciones realizadas por los hombres:

	Bonito	Feo
<i>Mujer</i>	98	152
<i>Hombre</i>	105	145

Se observa que aparentemente no existe un sesgo entre el genero de la persona de la foto y la percepción que se tiene sobre esta en los resultados generales obtenidos.

Resultados Hombres vs Resultados Mujeres

	Bonito	Feo
<i>Mujer</i>	41	84
<i>Hombre</i>	48	77

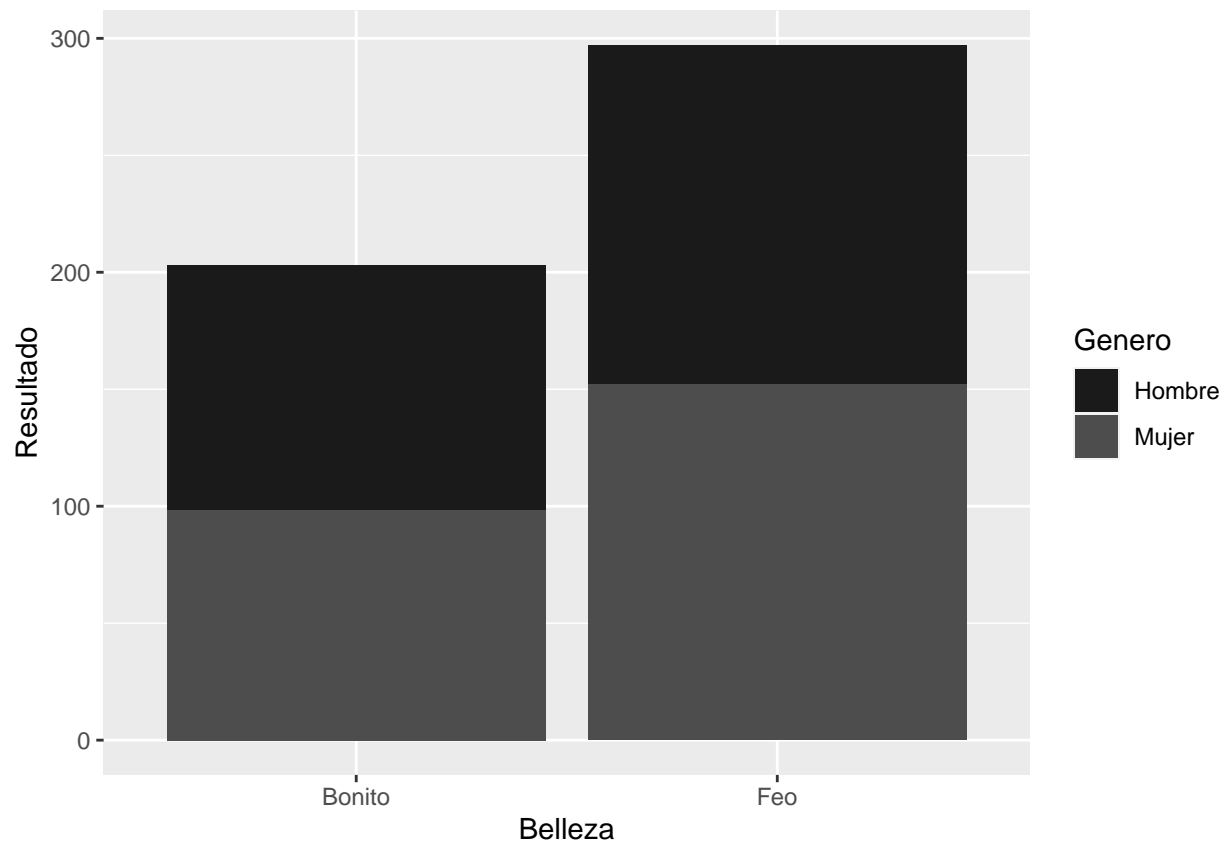
	Bonito	Feo
<i>Mujer</i>	57	68
<i>Hombre</i>	57	68

Se observa que para los resultados obtenidos en ambos sexo aparentemente no existe un sesgo entre el genero de la persona de la foto y la percepción que se tiene sobre esta.

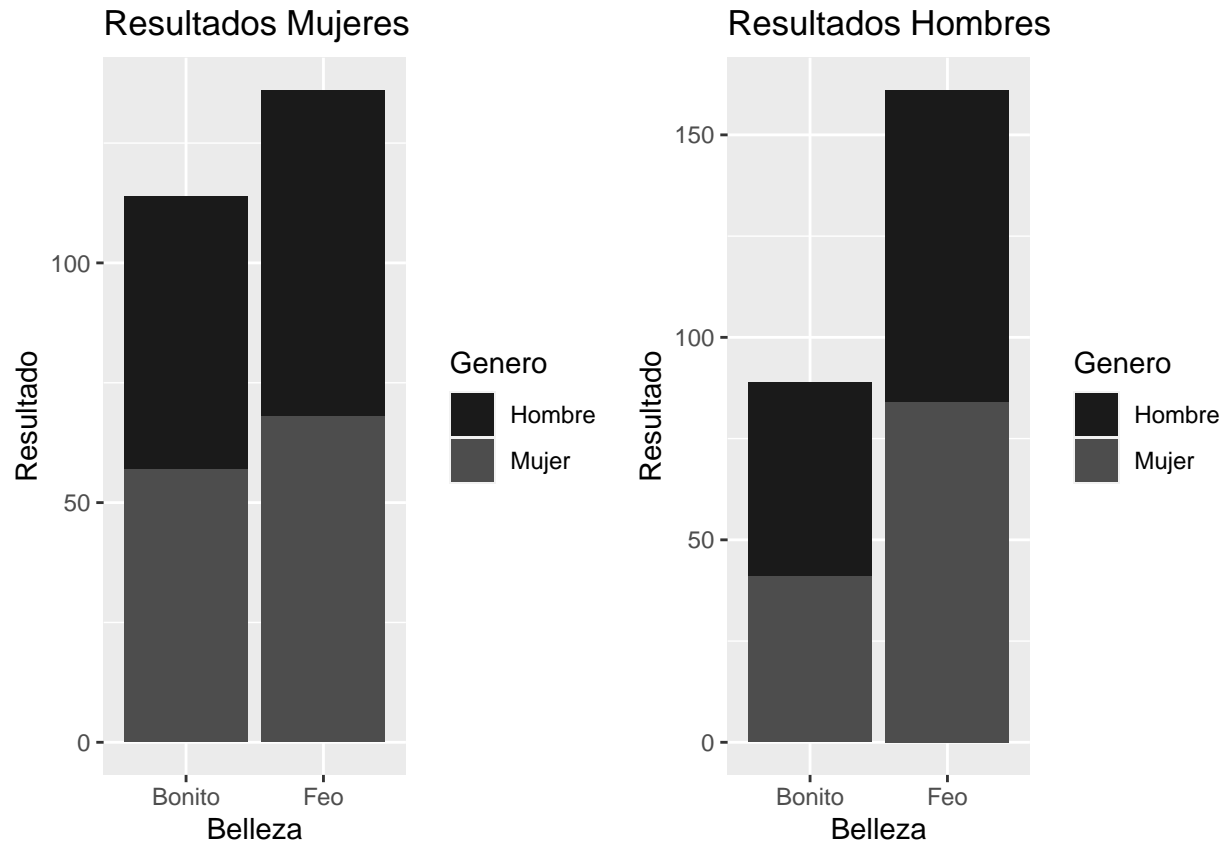
A partir de la información obtenida se plantea lo siguiente:

- Realizar un análisis descriptivo acerca de los resultados obtenidos en donde se intente observar si existe una posible asociación entre el Sexo de las personas que están en las imagenes y la percepción que tuvieron las personas encuestadas respecto a si esas personas son bonitas o feas.
- Realizar la prueba exacta de Fisher sobre los resultados generales y segregados por sexo para determinar si es posible que estos cambien respecto al sexo de la persona a la que se entreviste.

Se realizan gráficas generales en donde se intenta observar posibles asociaciones entre variables:



A partir de la gráfica general no se puede observar alguna inclinación a la percepción de belleza dependiendo del genero.



A partir de las gráficas se puede observar que graficamente no se observa posible disparidad entre el genero y la belleza percibida tanto de manera general como de manera segregada en los encuestados.

A continuación se plantean las correspondientes pruebas exactas de Fisher:

Prueba de Fisher para la tabla general

Bajo el esquema de la prueba de Fisher se plantea el siguiente juego de hipótesis:

H_0 : No existe isparidad entre el genero y la belleza percibida.

H_a : Existe evidencia de isparidad entre el genero y la belleza percibida.

Se rechaza H_0 sí $V_p < 0.05$

Fisher's Exact Test for Count Data

```
data:  tabla_general
p-value = 0.5848
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.6130576 1.2928945
sample estimates:
odds ratio
 0.8905592
```

Corroborando con la información gráfica a partir del resultado de la prueba no se puede concluir que exista

relación entre el género y la percepción que se tiene sobre la belleza de las personas en las fotos para las personas encuestadas.

Prueba de Fisher para las respuestas de las mujeres

Bajo el esquema de la prueba de Fisher se plantea el siguiente juego de hipótesis:

H_0 : No existe isparidad entre el genero y la belleza percibida.

H_a : Existe evidencia de isparidad entre el genero y la belleza percibida.

Se rechaza H_0 sí $V_p < 0.05$.

Fisher's Exact Test for Count Data

```
data:  tabla_mujeres
p-value = 1
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.5896521 1.6959152
sample estimates:
odds ratio
      1
```

Corroborando con la información gráfica a partir del resultado de la prueba no se puede concluir que exista relación entre el género y la percepción que se tiene sobre la belleza de las personas en las fotos para las mujeres encuestadas.

Prueba de Fisher para las respuestas de los hombres

Bajo el esquema de la prueba de Fisher se plantea el siguiente juego de hipótesis:

H_0 : No existe isparidad entre el genero y la belleza percibida.

H_a : Existe evidencia de isparidad entre el genero y la belleza percibida.

Se rechaza H_0 sí $V_p < 0.05$.

Fisher's Exact Test for Count Data

```
data:  tabla_hombres
p-value = 0.4281
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.4504582 1.3593071
sample estimates:
odds ratio
 0.7837607
```

Corroborando con la información gráfica a partir del resultado de la prueba no se puede concluir que exista relación entre el género y la percepción que se tiene sobre la belleza de las personas en las fotos para los hombres encuestados.

A partir de los resultados obtenidos se pueden realizar las siguientes conclusiones:

- Según los resultados en las pruebas realizadas no se tiene evidencia acerca de la existencia de algún sesgo de belleza asociado al sexo de las personas encontradas en las fotografías tanto para los resultados generales como para los resultados segregados por el sexo de las personas encuestadas.
- Tener en cuenta de que para obtener los datos utilizados en las pruebas no se tuvo de referente ningún esquema de muestreo para elegir a las personas, pero sí se tuvo en cuenta que por la naturaleza del experimento la clasificación de las imágenes sería una variable binomial.