

Ejercitación Unidad 2 - Parte 2

1. a. Tabla de contingencia

Cuadro 1: Tabla de contingencia para las variables
hábito de fumar y sobrevivida a 20 años

	Fumadora	No Fumadora	Total
No Sobrevive	139	230	369
Sobrevive	443	502	945
Total	582	732	1314

Medidas de asociación

Diferencia de proporciones condicionales:

$$\frac{139}{369} - \frac{443}{945} = -0,0920893$$

La proporción de fumadoras es menor entre los no sobrevivientes que entre los sobrevivientes

$$\frac{230}{369} - \frac{502}{945} = 0,0920893$$

La proporción de no fumadoras es mayor entre los no sobrevivientes que entre los sobrevivientes

Coefficiente de proporciones condicionales (o riesgo relativo)

$$RR = \frac{139}{369} / \frac{443}{945} = 0,8035567$$

La chance de ser fumadora es casi un 20 % menor entre los no sobrevivientes que entre los sobrevivientes.

$$RR = \frac{230}{369} / \frac{502}{945} = 1,173355$$

La chance de ser no fumadora es casi un 17 % mayor entre los no sobrevivientes que entre los sobrevivientes.

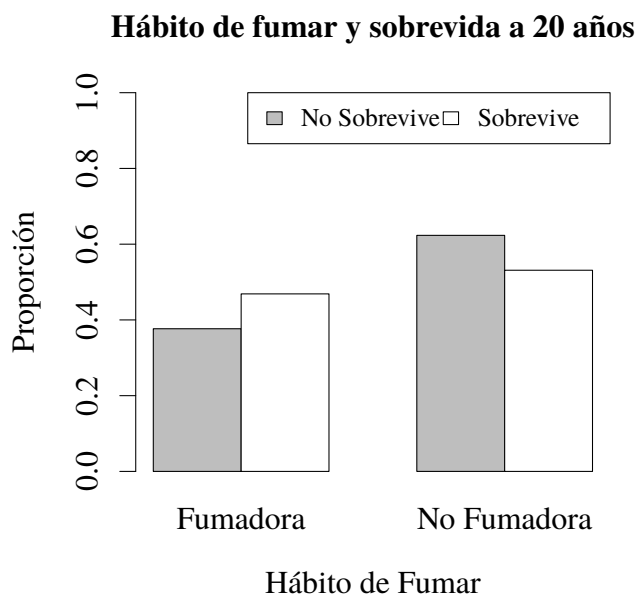


Figura 1: Proporción de sobrevida en fumadoras y no fumadoras

Test de independencia

Chi-cuadrado = 9.1209

p-value = 0.002527

El p-value es menor que $\alpha = 0.05$, por lo tanto se rechaza la hipótesis nula de independencia. Con lo cual el hábito de fumar está significativamente asociado a la sobrevida a 20 años.

Como se puede observar en la figura 1, el porcentaje de sobrevida es mayor en los fumadores, sumado a los valores observados en las medidas de asociación, los resultados resultan contrarios a lo esperado.

- b. A raíz de los resultados contradictorios que arroja el análisis anterior, los investigadores decidieron tener en cuenta en el análisis una variable adicional.

Cuadro 2: Estudio con variable Edad de las personas

Fuma	18-34 años		35-54 años		55-64 años		65 años o más	
	Sobrevive	No Sobrevive	Sobrevive	No Sobrevive	Sobrevive	No Sobrevive	Sobrevive	No Sobrevive
Si	174	5	198	41	64	51	7	42
No	213	6	180	19	81	40	28	165

Edad de 18 a 34 años

Cuadro 3: Tabla de contingencia para las variables
hábito de fumar y sobrevivida a 20 años

	Fumadora	No Fumadora	Total
No Sobrevive	5	6	11
Sobrevive	174	213	387
Total	179	219	398

Medidas de asociación

Diferencia de proporciones condicionales:

$$\frac{5}{11} - \frac{174}{387} = 0,004933051$$

La proporción de fumadoras es mayor entre los no sobrevivientes que entre los sobrevivientes

$$\frac{6}{11} - \frac{213}{387} = -0,004933051$$

La proporción de no fumadoras es menor entre los no sobrevivientes que entre los sobrevivientes

Coficiente de proporciones condicionales (o riesgo relativo)

$$RR = \frac{5}{11} / \frac{174}{387} = 1,010972$$

La chance de ser fumadora es casi igual entre sobrevivientes y no sobrevivientes.

$$RR = \frac{6}{11} / \frac{213}{387} = 0,9910371$$

La chance de ser no fumadora es casi igual entre sobrevivientes y no sobrevivientes.

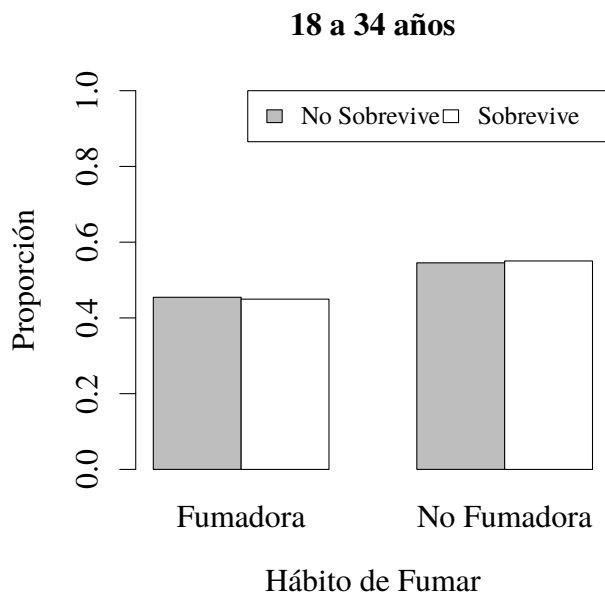


Figura 2: Proporción de sobrevivencia en fumadoras y no fumadoras

Chi-cuadrado = 0.0015576

p-value = 0.9685

El p-value es menor a $\alpha = 0,05$ por lo que se acepta la hipótesis nula de independencia. Además las medidas de asociación confirman esta independencia.

Edad de 35 a 54 años

Cuadro 4: Tabla de contingencia para las variables hábito de fumar y sobrevivencia a 20 años

	Fumadora	No Fumadora	Total
No Sobrevive	41	19	60
Sobrevive	198	180	378
Total	239	199	438

Medidas de asociación

Diferencia de proporciones condicionales:

$$\frac{41}{60} - \frac{198}{378} = 0,1595238$$

La proporción de fumadoras es mayor entre los no sobrevivientes que entre los sobrevivientes

$$\frac{19}{60} - \frac{180}{378} = -0,1595238$$

La proporción de no fumadoras es menor entre los no sobrevivientes que entre los sobrevivientes

Coefficiente de proporciones condicionales (o riesgo relativo)

$$RR = \frac{41/60}{198/378} = 1,304545$$

La chance de ser fumadora es casi un 30 % mayor entre los no sobrevivientes que entre los sobrevivientes.

$$RR = \frac{19/60}{180/378} = 0,665$$

La chance de ser no fumadora es un 33 % menor entre los no sobrevivientes que entre los sobrevivientes.

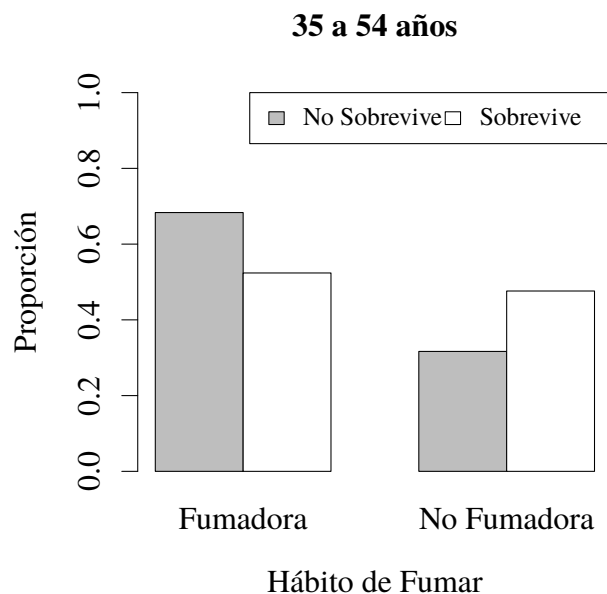


Figura 3: Proporción de sobrevida en fumadoras y no fumadoras

Chi-cuadrado = 5.3152

p-value = 0.02114

El p-value es menor al $\alpha = 0,05$ con lo cual se rechaza la hipótesis nula de independencia. Observando las medidas de asociación se encuentra que entre los no sobrevivientes, la chance de ser fumador es más alta. En la figura 3. esto se puede visualizar más claramente.

Edad de 55 a 64 años

Medidas de asociación

Diferencia de proporciones condicionales:

Cuadro 5: Tabla de contingencia para las variables
hábito de fumar y sobrevivida a 20 años

	Fumadora	No Fumadora	Total
No Sobrevive	51	40	91
Sobrevive	64	81	145
Total	115	121	236

$$\frac{51}{91} - \frac{64}{145} = 0,1190603$$

La proporción de fumadoras es mayor entre los no sobrevivientes que entre los sobrevivientes

$$\frac{40}{91} - \frac{81}{145} = -0,1190603$$

La proporción de no fumadoras es menor entre los no sobrevivientes que entre los sobrevivientes

Coefficiente de proporciones condicionales (o riesgo relativo)

$$RR = \frac{51}{91} / \frac{64}{145} = 1,269746$$

La chance de ser fumadora es un 26 % mayor entre los no sobrevivientes que entre los sobrevivientes.

$$RR = \frac{40}{91} / \frac{81}{145} = 0,7868675$$

La chance de ser no fumadora es casi un 22 % menor entre los no sobrevivientes que entre los sobrevivientes.

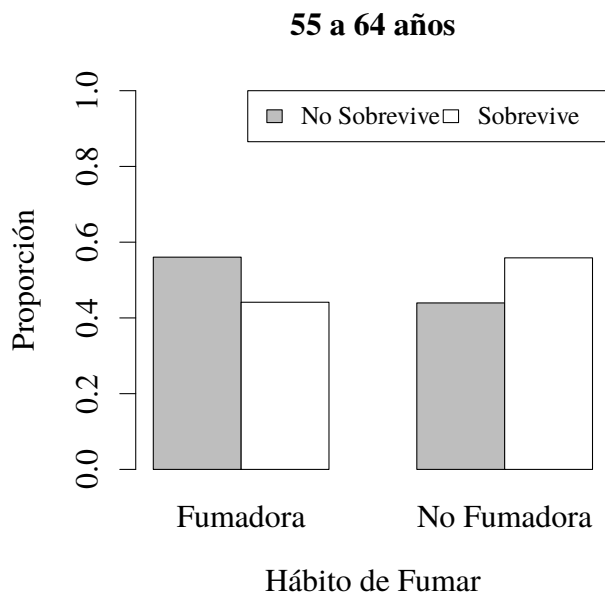


Figura 4: Proporción de sobrevivencia en fumadoras y no fumadoras

Chi-cuadrado = 3.1723

p-value = 0.0749

El p-value es mayor al $\alpha = 0,05$ con lo cual se acepta la hipótesis nula de independencia. Sin embargo, al observar las medidas de asociación, se encuentra que entre los no sobrevivientes, la chance de ser fumador sigue siendo más alta. En la figura 4. se puede visualizar este resultado.

Edad de 65 años o más

Cuadro 6: Tabla de contingencia para las variables hábito de fumar y sobrevivencia a 20 años

	Fumadora	No Fumadora	Total
No Sobrevive	42	165	207
Sobrevive	7	28	35
Total	49	193	242

Medidas de asociación

Diferencia de proporciones condicionales:

$$\frac{42}{207} - \frac{7}{35} = 0,002898551$$

La proporción de fumadoras es mayor entre los no sobrevivientes que entre los sobrevivientes

$$\frac{165}{207} - \frac{28}{35} = -0,002898551$$

La proporción de no fumadoras es menor entre los no sobrevivientes que entre los sobrevivientes

Coefficiente de proporciones condicionales (o riesgo relativo)

$$RR = \frac{42}{207} / \frac{7}{35} = 1,014493$$

La chance de ser fumadora es casi igual entre sobrevivientes y no sobrevivientes.

$$RR = \frac{165}{207} / \frac{28}{35} = 0,9963768$$

La chance de ser fumadora es casi igual entre sobrevivientes y no sobrevivientes.

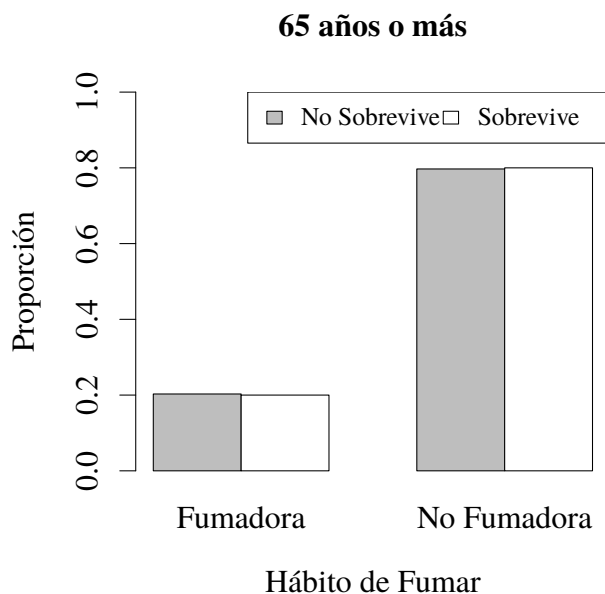


Figura 5: Proporción de sobrevida en fumadoras y no fumadoras

Chi-cuadrado = 0.0015576

p-value = 0.9685

El p-value es mayor al $\alpha = 0,05$ con lo cual se acepta la hipótesis nula de independencia. Sin embargo, al observar las medidas de asociación, se puede concluir que la sobrevida y el hábito de fumar no están relacionados. Eso mismo se puede concluir observando la figura 5.

Conclusión: Observando las medidas de asociación y los gráficos, es posible concluir que la relación entre el hábito de fumar y sobrevivir es muy alta en las personas entre 35 a 64 años de edad. Especialmente en el rango de edad entre 35 a 54 años, las personas fumadoras tienen

más chance de no sobrevivir.

2. Esquema Ahorro - Inversión - Financiamiento. Administración Pública Nacional. Valores mensuales. En millones de pesos. Metodología 2017
Fuente de datos: datos.gob.ar

Cantidad de registros: 51 observaciones

Datos mensuales desde 2015-01-01 al 2019-03-01.

Cuadro 7: Resumen descriptivo de cada variable

ing_corr_2017	gtos_corr_2017
Min. : 69844	Min. : 71315
1st Qu.:106799	1st Qu.:108740
Median :144242	Median :158267
Mean :144937	Mean :164838
3rd Qu.:179648	3rd Qu.:193166
Max. :252500	Max. :360150

Diagrama de dispersión

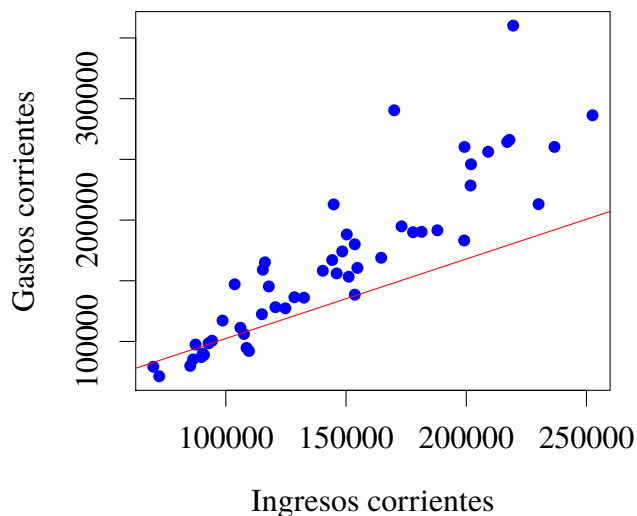


Figura 6: Diagrama de dispersión de gastos e ingresos corrientes

Coefficiente de correlación de Pearson:

$r = 0,9059309$

Gráfico de dispersión - Rangos

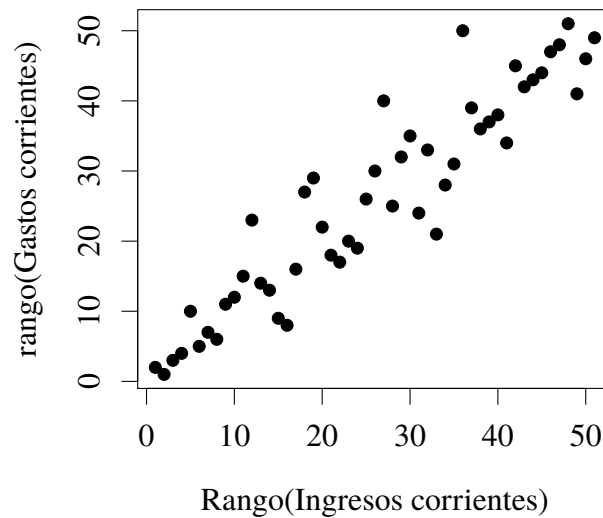


Figura 7: Diagrama de dispersión de rangos de gastos e ingresos corrientes

Coefficiente de correlación de Spearman:

$$r_s = 0,9377376$$

Conclusión: Los coeficientes de correlación de Pearson y Spearman son cercanos a 1, lo cual muestra una fuerte asociación lineal positiva entre las variables de Gastos corrientes e Ingresos corrientes.