

# Inferencia Estadística

## Tablas de contingencia

Edimer David Jaramillo - Bioestadística 1

Abril de 2019

# Tablas de contingencia

# Introducción

- Ideal para analizar variables categóricas.
- Los datos se presentan como frecuencias de observaciones.
- De gran utilidad para verificar si los valores de dos factores son dependientes o independientes.
- El estadístico  $\chi^2$  es útil para el análisis de variables categóricas

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Donde:

- **O**: frecuencia observada.
- **E**: frecuencia esperada.

# Ejemplo

# Enunciado

Se dispone de 300 animales de laboratorio y se decide tratar a 200 con una vacuna experimental y dejar 100 como controles. Después de tratar al primer lote se expone a los 300 al contagio de la enfermedad en estudio. El recuento final, después de un período experimental adecuado, fue:

|             | Enfermos | Sanos | Total |
|-------------|----------|-------|-------|
| Tratados    | 56       | 144   | 200   |
| No Tratados | 71       | 29    | 100   |
| Total       | 127      | 173   | 300   |

$H_0$  : los dos factores o variables categóricas son independientes

$H_0$  : los dos factores o variables categóricas no son independientes

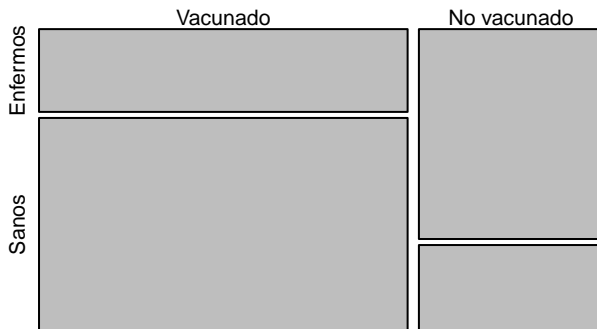
# Solución con R (1/3)

```
tabla <- matrix(c(56, 144, 71, 29), nrow = 2, ncol = 2,  
               byrow = TRUE)  
colnames(tabla) <- c("Enfermos", "Sanos")  
rownames(tabla) <- c("Vacunado", "No vacunado")  
tabla  
##           Enfermos Sanos  
## Vacunado           56   144  
## No vacunado        71    29
```

# Solución con R (2/3)

```
mosaicplot(tabla, main = "Gráfico de mosaico")
```

## Gráfico de mosaico





## Solución con R (3/3)

- Función `chisq.test()`
  - `x`: tabla de datos.

```
prueba_chi2 <- chisq.test(tabla, correct = FALSE)
prueba_chi2
##
## Pearson's Chi-squared test
##
## data:  tabla
## X-squared = 50.494, df = 1, p-value = 1.195e-12
```

```
prueba_chi2$observed
```

| ##             | Enfermos | Sanos |
|----------------|----------|-------|
| ## Vacunado    | 56       | 144   |
| ## No vacunado | 71       | 29    |

# Frecuencias esperadas

prueba\_chi2\$expected

| ##             | Enfermos | Sanos     |
|----------------|----------|-----------|
| ## Vacunado    | 84.66667 | 115.33333 |
| ## No vacunado | 42.33333 | 57.66667  |

¿De dónde salen estos valores esperados?

$$F_{\text{esperada-vacunados}} = \frac{200 * 127}{300} = 84.667$$

$$F_{\text{esperada-sanos-vacunados}} = \frac{200 * 173}{300} = 115.33$$