

# Metodo tablas de contingencia

Grupo Sierra

2024-05-22

# Tabla de contingencias

Una tabla de contingencia es una herramienta estadística que se utiliza para mostrar y analizar la relación entre dos o más variables categóricas. Organiza los datos en una matriz donde las filas representan una variable y las columnas otra, permitiendo observar la frecuencia de las combinaciones posibles de estas variables. Es útil para identificar patrones y relaciones, y se puede utilizar para realizar pruebas estadísticas como el chi-cuadrado. (El chi-cuadrado es una prueba estadística que se utiliza para evaluar si existe una asociación significativa entre dos variables categóricas)

```
extrovertidos <- c(60, 40) # Personas extrovertidas que ganan > $50,000 y <= $50,000
introvertidos <- c(40, 60) # Personas introvertidas que ganan > $50,000 y <= $50,000
tabla_contingencia <- matrix(c(extrovertidos, introvertidos), nrow=2, byrow=TRUE)
colnames(tabla_contingencia) <- c("Ganan > $50,000", "Ganan <= $50,000")
rownames(tabla_contingencia) <- c("Extrovertidos", "Introvertidos")
print(tabla_contingencia)
```

```
##           Ganan > $50,000 Ganan <= $50,000
## Extrovertidos           60           40
## Introvertidos           40           60
```

# Cuando usar tablas de contingencias

Algunas situaciones y contextos en los que se utilizan las tablas de contingencia:

Análisis de independencia

Estudios de asociación

Investigación en ciencias sociales

Marketing y análisis de consumidores

# Ventajas y desventajas

## VENTAJAS

**Simplicidad y claridad:** Las tablas de contingencia son fáciles de entender y permiten visualizar rápidamente la relación entre dos o más variables categóricas.

**Versatilidad:** Pueden ser utilizadas en una amplia variedad de campos, como la biomedicina, las ciencias sociales, el marketing, y más.

**Análisis de independencia:** Facilitan la realización de pruebas de independencia como la chi-cuadrado, que ayuda a determinar si existe una relación significativa entre las variables.

## DESVENTAJAS

**Limitación a datos categóricos:** Solo pueden utilizarse con variables categóricas, no con variables continuas, lo que limita su aplicabilidad.

**Tamaño limitado:** Las tablas grandes con muchas categorías pueden volverse complicadas y difíciles de interpretar. Además, pueden ser poco prácticas si hay muchas categorías con frecuencias bajas.

# Grafico de puntos

```
frecuencia <- c(60, 40, 40, 60)
personalidad_ingresos <- c("Extrovertidos > $50,000", "Extrovertidos <= $50,000",
    "Introvertidos > $50,000", "Introvertidos <= $50,000")
dotchart(frecuencia, labels = personalidad_ingresos,
    pt.cex = 2, pch = c(8, 21), color = c("blue", "red"),
    lcolor = "black", xlab = "Número de personas",
    main = "Diagrama de puntos de ingresos por tipo de personalidad")
```

# Breve explicacion de como hacer el codigo

Primero hay que preparar los datos para el diagrama de puntos: Se combinan las frecuencias de personas extrovertidas e introvertidas que ganan más y menos de \$50,000 en un solo vector frecuencia. El vector `personalidad_ingresos` contiene las etiquetas correspondientes.

Crear el dotchart: Se utiliza la función `dotchart` con los parámetros `pt.cex`, `pch`, `color`, `lcolor`, y `xlab` para personalizar el gráfico:

`pt.cex = 2`: Tamaño de los puntos. `pch = c(8, 21)`: Tipo de símbolo para los puntos. `color = c("blue", "red")`: Colores para los puntos (azul para extrovertidos y rojo para introvertidos). `lcolor = "black"`: Color de las líneas. `xlab = "Número de personas"`: Etiqueta del eje x. `main = "Diagrama de puntos de ingresos por tipo de personalidad"`: Título del gráfico