

# **Práctica 3A: Probabilidad y Estadística Descriptiva**

Alumno: Javier Jhairt Lopez Rojas  
Materia: Probabilidad y Estadística

# Problema 1

**Postulado:** Lanzar tres veces una moneda honesta y calcular probabilidades.

## Solución

### (a) Espacio muestral

$$\Omega = \{AAA, AAS, ASA, ASS, SAA, SAS, SSA, SSS\}$$

### (b) Función de probabilidad

Sea  $X$  el número de águilas. Entonces:

$$X \sim Binomial(n = 3, p = 0,5)$$

$$P(X = k) = \binom{3}{k} (0,5)^k (0,5)^{3-k}$$

$$P(0) = \frac{1}{8}, \quad P(1) = \frac{3}{8}, \quad P(2) = \frac{3}{8}, \quad P(3) = \frac{1}{8}$$

### (c) Esperanza

$$E(X) = np = 3(0,5) = 1,5$$

### (d) Varianza

$$Var(X) = np(1 - p) = 3(0,5)(0,5) = 0,75$$

**Interpretación:** En promedio se esperan 1.5 águilas en tres lanzamientos.

## Problema 2

**Postulado:** Examen de 10 preguntas con probabilidad  $p = 0,25$ .

### Solución

$$X \sim Binomial(n = 10, p = 0,25)$$

(a) Exactamente 3 correctas

$$P(X = 3) = \binom{10}{3} (0,25)^3 (0,75)^7$$

(b) Al menos 2 correctas

$$P(X \geq 2) = 1 - P(X \leq 1)$$

(c) Ninguna correcta

$$P(X = 0) = (0,75)^{10}$$

**Interpretación:** Se modela con distribución binomial debido a ensayos independientes.

### Anexo: Código en R

```
n <- 10  
p <- 0.25  
dbinom(3, n, p)  
1 - pbinom(1, n, p)  
dbinom(0, n, p)
```

## Problema 3

**Postulado:** Indemnizaciones en miles de pesos.

### Solución

Se calcularon media, mediana, varianza, desviación estándar, percentiles 0.16 y 0.84, así como deciles 1, 5 y 9.

#### Interpretación:

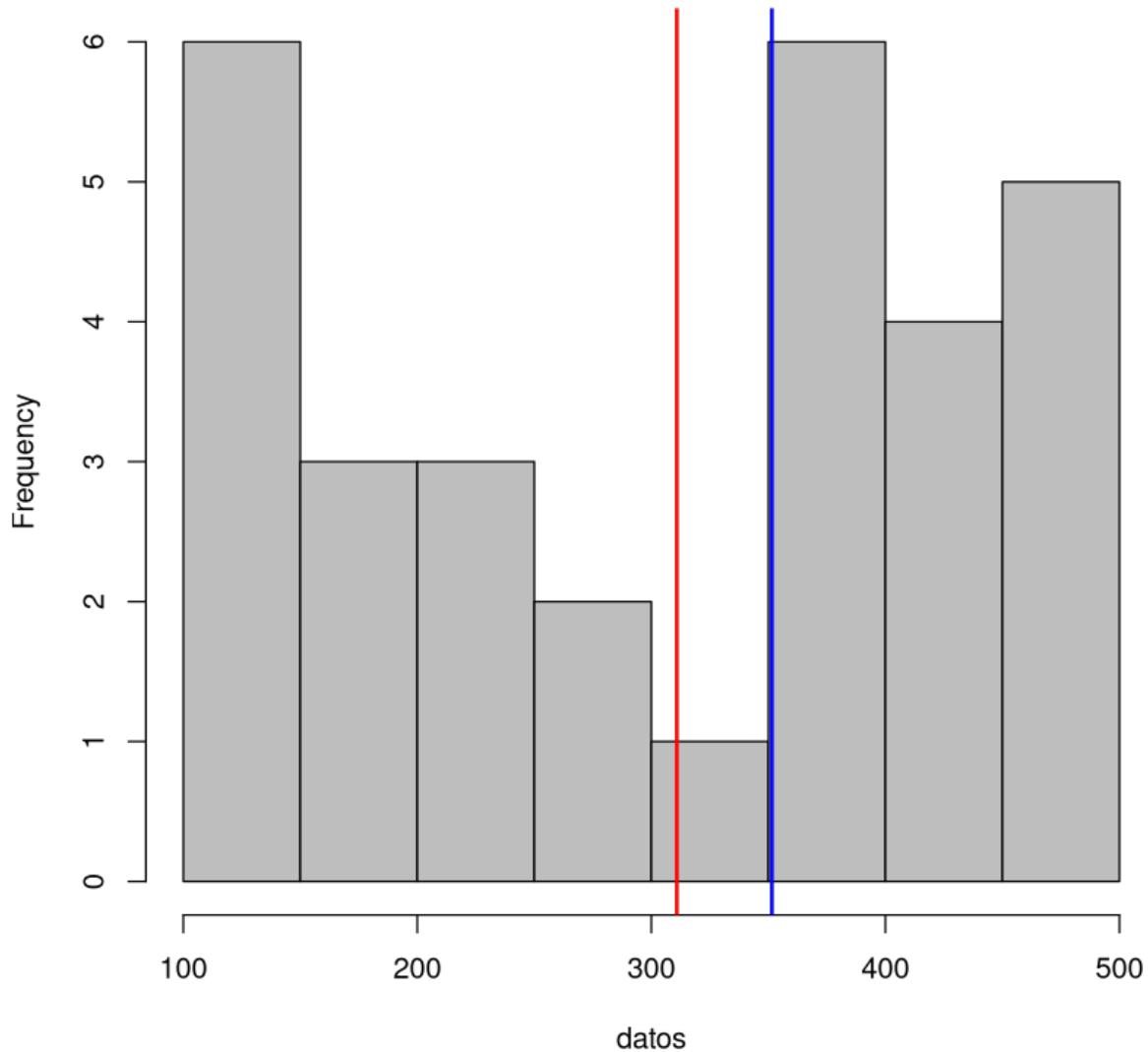
- La línea roja representa la **media**, y la línea azul la **mediana**.
- La media se encuentra ligeramente a la derecha de la mediana.
- Esto indica una **asimetría positiva**, causada por algunos valores altos en el conjunto de datos.
- La mayoría de las indemnizaciones se concentran en valores medios y bajos.
- Existen algunos montos elevados que influyen en el promedio.
- Debido a la influencia de valores altos, la **media no es completamente representativa**, y la mediana describe mejor el centro del conjunto.
- En general, los datos muestran dispersión moderada con presencia de valores extremos superiores.

### Anexo: Código en R

```
datos <- c(500,500,495,490,473,441,429,419,405,400,  
390,390,390,376,353,350,300,258,240,220,210,200,  
192,190,150,130,125,110,100,100)
```

```
mean(datos)  
median(datos)  
var(datos)  
sd(datos)  
quantile(datos, c(0.16,0.84))  
quantile(datos, c(0.1,0.5,0.9))  
  
hist(datos)  
abline(v=mean(datos), col="red")  
abline(v=median(datos), col="blue")
```

## Histogram of datos



## Problema 4

**Postulado:** Tabla de contingencia Sexo vs Estado Civil.

### Solución

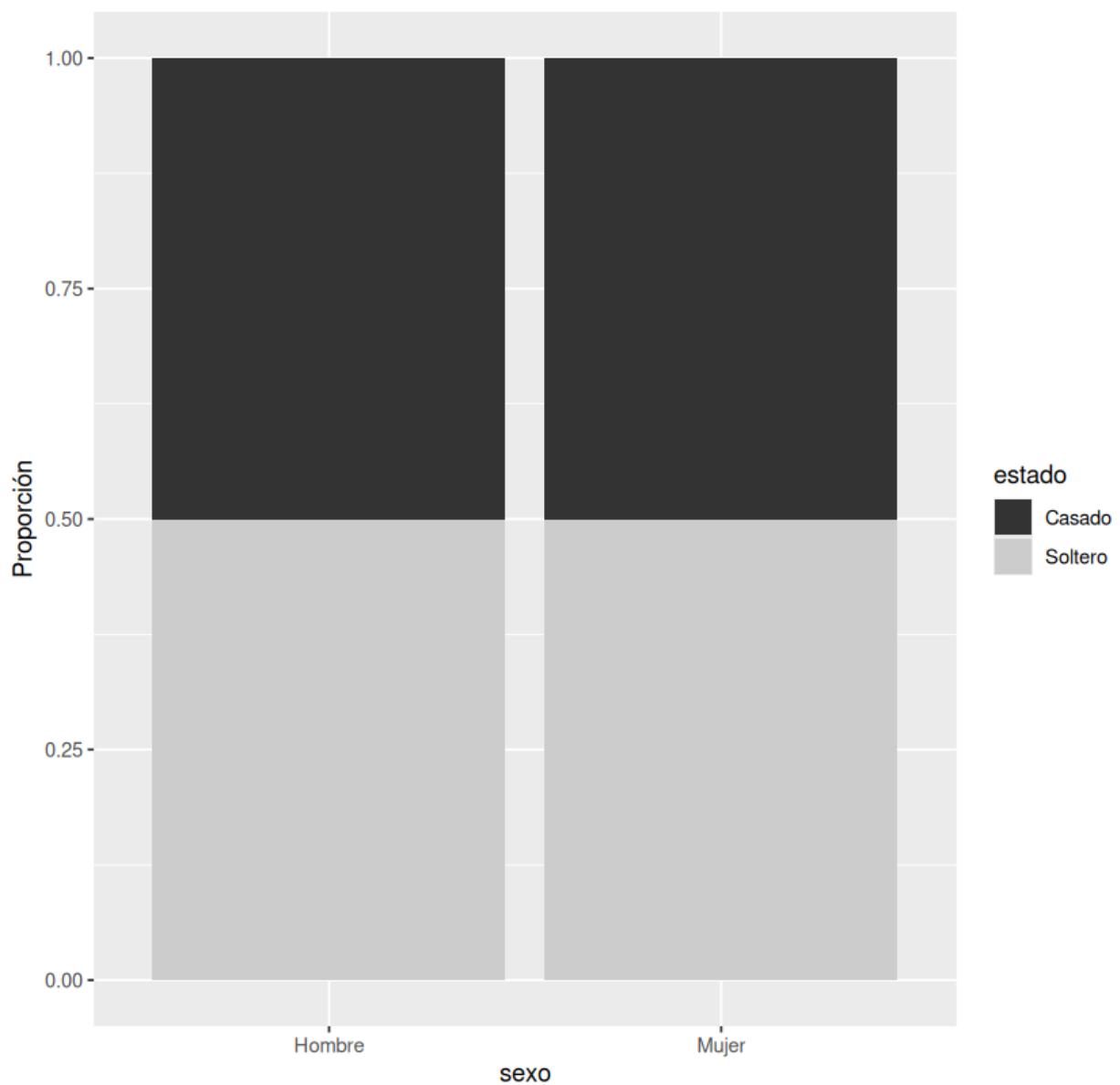
Se construyó tabla de contingencia, marginal y condicional.

**Interpretación:** En la tabla de contingencia se observa que tanto hombres como mujeres están distribuidos de manera equilibrada entre casados y solteros, ya que en ambos casos el 50 % pertenece a cada categoría. Esto indica que dentro de cada sexo no hay una diferencia en la proporción de estado civil. Además, aunque hay ligeramente más hombres que mujeres en la empresa (20 hombres y 16 mujeres), esa diferencia se refleja de manera proporcional en ambas categorías de estado civil. En general, no se aprecia una relación fuerte entre sexo y estado civil, por lo que las variables parecen comportarse de forma independiente en esta muestra.

### Anexo: Código en R

```
tabla <- table(sexo, estado)
margin.table(tabla,1)
margin.table(tabla,2)
prop.table(tabla,1)
prop.table(tabla,2)
```

Proporción de Estado Civil por Sexo



## Problema 5

**Postulado:** Datos USArrests.

### Solución

Se calcularon:

- Matriz varianza-covarianza
- Matriz de correlación
- Scatter Plot Matrix
- Parallel Coordinates Plot

**Interpretación:** En la matriz de dispersión se observa que Murder y Assault tienen una relación positiva fuerte, ya que cuando una variable aumenta, la otra también tiende a aumentar. Rape muestra una relación moderada con estas variables, mientras que UrbanPop no presenta una relación tan clara. En el gráfico de coordenadas paralelas se pueden notar líneas que siguen patrones similares, lo que indica que existen estados con niveles parecidos de criminalidad. En general, ambas gráficas permiten identificar visualmente grupos de estados con niveles altos, medios y bajos en las variables analizadas.

### Anexo: Código en R

```
data(USArrests)
var(USArrests)
cor(USArrests)
pairs(USArrests)
matplot(scale(USArrests), type="l")
```

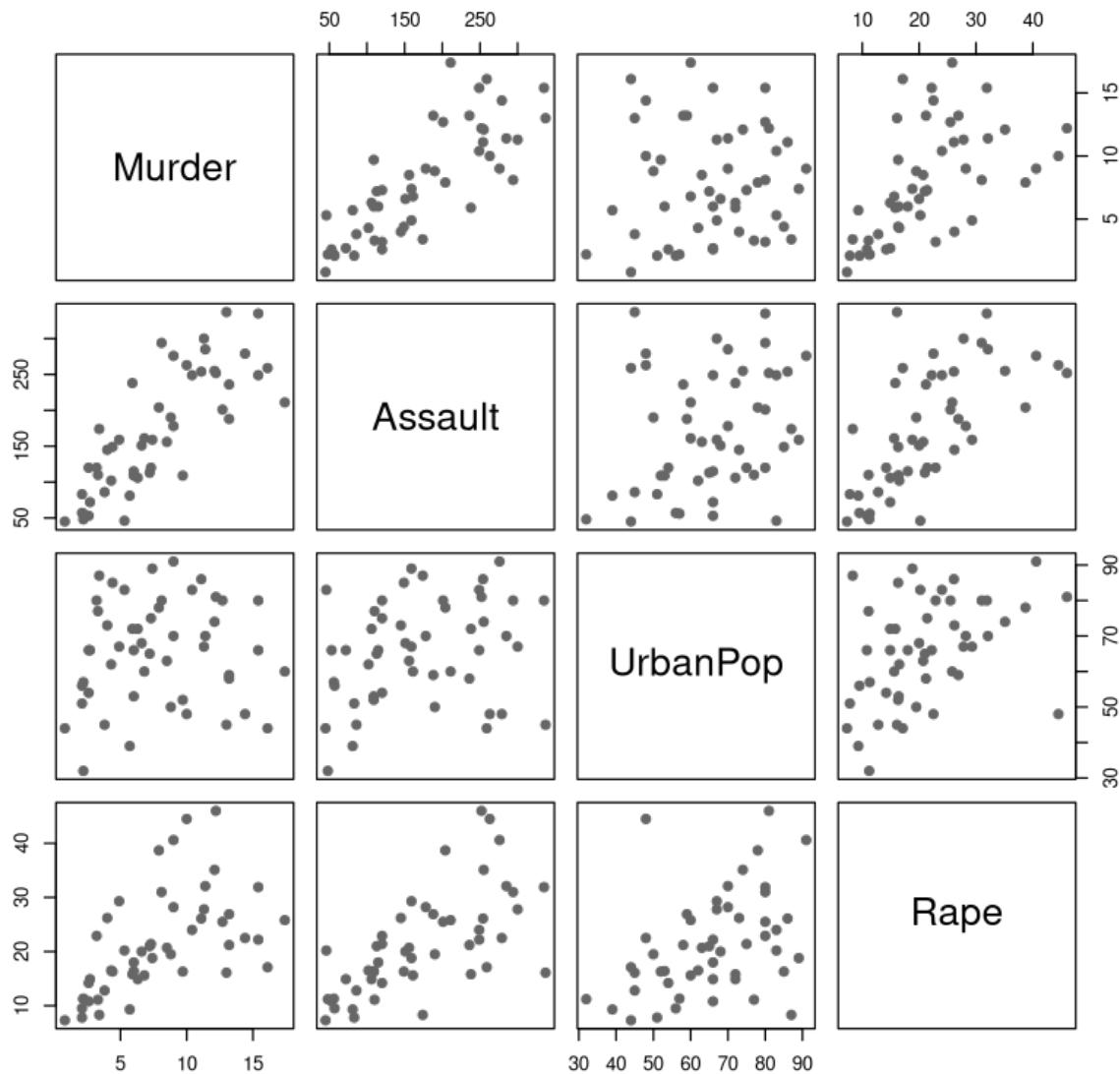
A matrix: 4 × 4 of type dbl

	<b>Murder</b>	<b>Assault</b>	<b>UrbanPop</b>	<b>Rape</b>
<b>Murder</b>	18.970465	291.0624	4.386204	22.99141
<b>Assault</b>	291.062367	6945.1657	312.275102	519.26906
<b>UrbanPop</b>	4.386204	312.2751	209.518776	55.76808
<b>Rape</b>	22.991412	519.2691	55.768082	87.72916

A matrix: 4 × 4 of type dbl

	<b>Murder</b>	<b>Assault</b>	<b>UrbanPop</b>	<b>Rape</b>
<b>Murder</b>	1.00000000	0.8018733	0.06957262	0.5635788
<b>Assault</b>	0.80187331	1.0000000	0.25887170	0.6652412
<b>UrbanPop</b>	0.06957262	0.2588717	1.00000000	0.4113412
<b>Rape</b>	0.56357883	0.6652412	0.41134124	1.0000000

### Scatter Plot Matrix - USArrests



## Parallel Coordinates Plot - USArrests

