HighwayMPG	n_i	f_i	N_i	F_i
$[20, 26)$			22	
$[26, 32)$			71	
$[32, 38)$			88	
$[38, 44)$			91	
$[44, 50]$			93	

Horsepower	n_i	f_i	N_i	F_i
$[54.8, 104)$				0.26
$[104, 153)$				0.59
$[153, 202)$				0.88
$[202, 251)$				0.95
$[251, 300]$				1.00

Conjunto de datos *cars93* del plugin. N=93

2.3 Principales características numéricas

La información suministrada por una tabla de distribución de frecuencias se puede resumir en un conjunto de medidas que la caracterizan y que se pueden clasificar en:

1. Medidas de posición: proporcionan valores que determinan posiciones dentro del conjunto de los datos. Las dividimos en medidas de tendencia central y de tendencia no central.

2. Medidas de dispersión: indican la desviación de los datos respecto de ciertas medidas de posición.

3. Medidas de forma: relacionadas con la representación gráfica de la distribución.

2.4 Medidas de posición de tendencia central

Media aritmética ($\overline{X}$): cociente entre la suma de todos los valores observados de la variable y el número total de observaciones:

$$\overline{X} = \frac{\sum_{i=1}^{k} X_i n_i}{N}$$

Para datos agrupados se toman como x_i las marcas de clase (representante de todos los datos del intervalo).

Propiedades

1. $\sum_{i=1}^{k}(x_i - \overline{x})n_i = 0.$
2. La media de las desviaciones al cuadrado de los valores de las variable respecto a una constante c cualquiera, se hace mínima cuando esa constante c es igual a la media aritmética. Es decir:

$$\min_c \sum_{i=1}^{k}(x_i - c)^2 n_i = \sum_{i=1}^{k}(x_i - \overline{x})^2 n_i \text{ (Teorema de König)}$$

3. Linealidad de la media: dados $a, b \in \mathbb{R}$ e $Y = a + bX$ se verifica:

$$\overline{y} = a + b\overline{x}$$

4. Media en subpoblaciones: Si el total de datos se estratifica en L grupos distintos, la media aritmética del total es una media aritmé tica de las distintas medias de los estratos ponderadas por el número de observaciones que tienen los mismos:

$$\overline{x} = \frac{\overline{x}_1 N_1 + \overline{x}_2 N_2 + \ldots + \overline{x}_L N_L}{N_1 + N_2 + \ldots + N_L}$$

5. $\min(x_i) \leq \overline{X} \leq \max(x_i)$

NOTA: A veces se introducen unos coeficientes de ponderación o pesos denominados w_i que son distintos de n_i con lo que la +media ponderada+ es:

$$\overline{x}_w = \frac{\sum_{i=1}^{k} w_i x_i}{\sum_i w_i}$$

Ejemplo 2.2. Calcula la media aritmética de los 10 primeros números naturales. *X= 10 primeros números enteros.* $x_1 = 1, x_2 = 2, \ldots, x_{10} = 10$, frecuencias $n_i = 1$, $i = 1, 2, \ldots, 10$, por lo tanto $\bar{x} = \frac{1+2+\ldots+10}{10} = \frac{55}{10} = 5.5$

Ejemplo 2.3. Supón ahora que la ponderación de cada valor es inversamente proporcional a su valor. $x_1 = 1, x_2 = 2, \ldots, x_{10} = 10$, con pesos $w_i = \frac{1}{i}$, $i = 1, 2, \ldots, 10$, por lo tanto $\bar{x}_w = \frac{1*\frac{1}{1}+2*\frac{1}{2}+\ldots+10*\frac{1}{10}}{\frac{1}{1}+*\frac{1}{2}+\ldots+*\frac{1}{10}} = \frac{10}{2.928968} = 3.414172$

Ejemplo 2.4. Haz los cálculos anteriores con R. Genera una función para calcular media aritmética y media ponderada.

Media geométrica: la raíz N-ésima del producto de los N valores de la distribución elevados al número de veces que se repite cada uno de ellos:

$$g = \sqrt[N]{x_1^{n_1} x_2^{n_2} \ldots x_k^{n_k}}$$

Podemos expresar la media geométrica como una media aritmética teniendo en cuenta que el logaritmo de la media geométrica es igual a la media aritmética de los logaritmos de los valores de la variable:

$$\log(g) = \frac{\sum_{i=1}^{k} \log(x_i)n_i}{N}$$

El empleo más frecuente de la media geométrica es el de promediar porcentajes, tasas, números índices, etc.

Mediana (Me): dada una distribución de frecuencias $\{(x_i, n_i)\}_{i=1}^{k}$ con valores ordenados de menor a mayor, llamamos mediana, *Me*, al valor de la variable que deja a su izquierda la misma frecuencia que a su derecha. Es decir, serían el valor o valores de la variable que son mayores o iguales que la mitad de los datos y menore o iguales a la otra mitad de los datos.

Cálculo de la mediana:

1. Distribuciones no agrupadas: Se observa cuál es la primera frecuencia absoluta acumulada N_i que supera o iguala a $N/2$ distinguiéndose dos casos:
   - Si $N_i > N/2$, entonces $Me = x_i$.
   - Si $N_i = N/2$, entonces $Me = \dfrac{x_i + x_{i+1}}{2}$.
2. Distribuciones agrupadas: el cálculo se realiza en dos etapas:
   - Detectar el *intervalo mediano*: $[L_{i-1}, L_i)$ tal que $N_i \geq \dfrac{N}{2}$.
   - Cálculo de la mediana mediante semejanza de triángulos:

$$Me = L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} a_i = L_{i-1} + \frac{\frac{1}{2} - F_{i-1}}{f_i} a_i$$

siendo L_{i-1}: el extremo inferior del intervalo mediano.

N_{i-1}: la frecuencia absoluta acumulada del intervalo anterior al mediano.

n_i: la frecuencia absoluta del intervalo mediano.

a_i: la amplitud del intervalo mediano.

Ejemplo 2.5. Calcula la mediana de la siguiente tabla de frecuencias:

x_i	n_i	N_i
1	1	1
2	1	2
3	1	3
4	1	4
	$N = 4$	

Variable discreta y como $N_2 = \frac{4}{2}$, la mediana es $\dfrac{x_i + x_{i+1}}{2} = 2.5$

Ejemplo 2.6. Calcula la mediana de la siguiente tabla de frecuencias:

x_i	n_i	N_i
1	10	10
2	5	15
3	15	30
4	10	40
	$N = 40$	

Variable discreta y como $N_3 = 30 > \frac{40}{2} = 20$, la mediana es 3

Ejemplo 2.7. Calcula la mediana de los 10 primeros números naturales. X= *10 primeros números enteros.* $x_1 = 1, x_2 = 2, \ldots, x_{10} = 10$, frecuencias $n_i = 1$, $i = 1, 2, \ldots, 10$, por lo tanto la variable es discreta (no agrupada). El vector de frecuencias acumuladas es $N_1 = 1, N_2 = 2, \ldots, N_{10} = 10$, el primer valor que es mayor o igual que $\frac{10}{2} = 5$ es $x_5 = 5$, como $N_5 = \frac{10}{2}$, entonces la mediana es $M_e = \frac{x_5+x_6}{2} = 5.5$. Coincide con la media.

Ejemplo 2.8. Supón ahora que la ponderación de cada valor es inversamente proporcional a su valor. $x_1 = 1, x_2 = 2, \ldots, x_{10} = 10$, con pesos $w_i = \frac{1}{i}$, $i = 1, 2, \ldots, 10$, El vector de frecuencias acumuladas es $(N_1 = 1, N_2 = 1.5, \ldots, N_{10} = 2.928968) = ((1, 1.5, 1.833, 2.0833, 2.2833, 2.45, 2.593, 2.7178, 2.828, 2.929)$, el primer valor que es mayor o igual que $\frac{2.928968}{2} = 1.464484$ es $x_2 = 1.5$, como es estrictamente mayor, entonces la mediana es $M_e = x_2 = 2$

Ejemplo 2.9. Calcula la mediana de la tabla agrupada del ejercicio de la pág. 4. El vector de frecuencias acumuladas es $(N_1 = 10, N_2 = 42, N_3 = 92, N_4 = 100)$, el primer índice que es mayor o igual que $\frac{100}{2} = 50$ es $i = 3$,el intervalo es $(28, 32]$, como es estrictamente mayor, entonces la mediana es

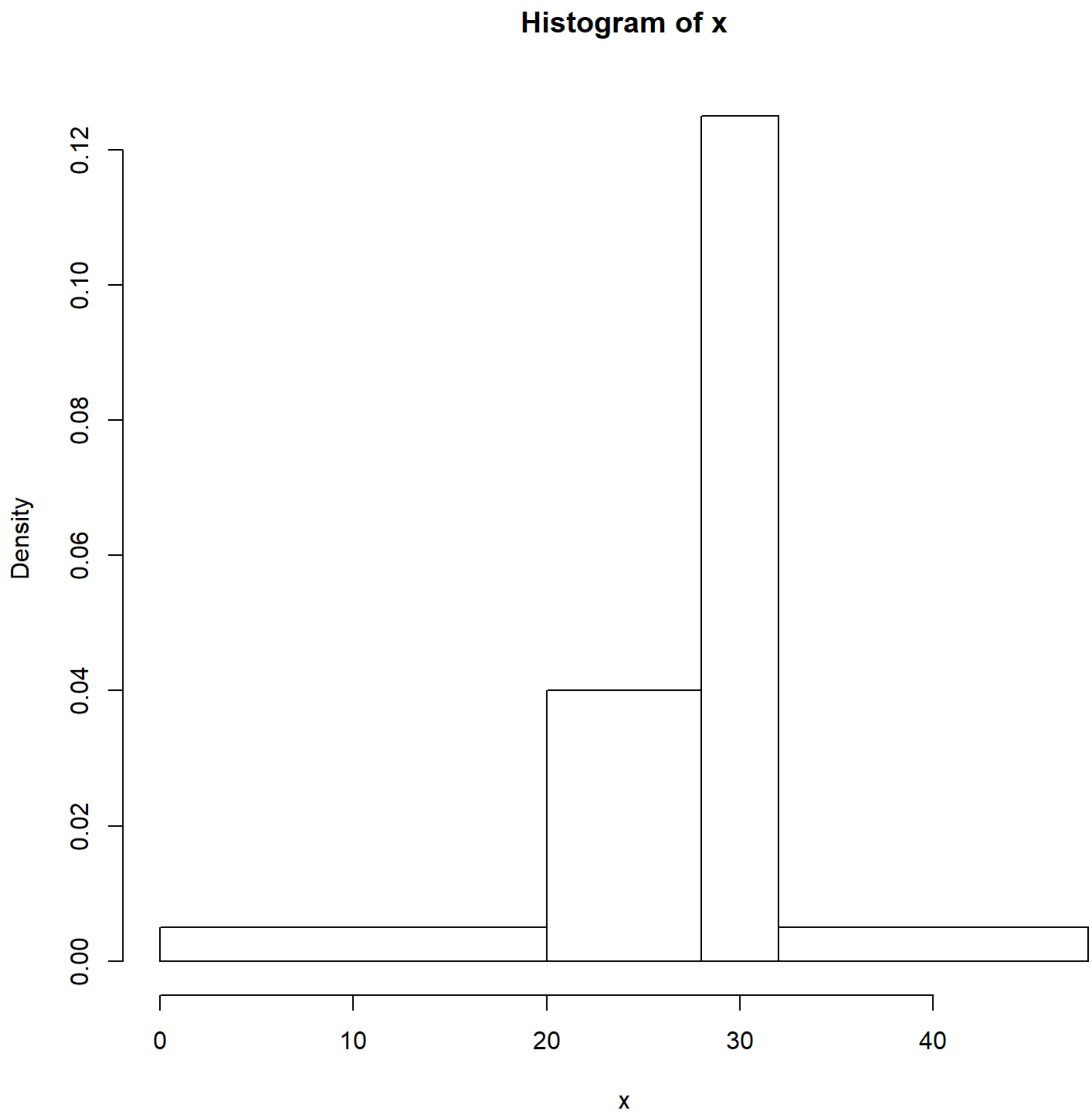
$$M_e = 28 + \frac{50 - 42}{50} * 4 = 28 + \frac{32}{50} = 28 + 0.64 = 28.64$$

Usando el paquete de R *actuar*

```
> ### Calculo con R
>   #install.packages("actuar")
>   require(actuar)          # instalar el paquete actuar previamente!
>          Li <- c(0, 20, 28, 32, 48)                  # limites de los intervalos
>         ni <- c(10,32,50,8)
>         x <- grouped.data(Group = Li, Frequency = ni)
>   #        x
>         quantile(x,probs=0.5)                # función quantile.grouped.data()
```

```
  50%
28.64
```

```
>         # codigo actuar:::quantile.grouped.data
>         hist(x) # histograma para datos agrupados --> función hist.grouped.data()
```


Histogram of x
Density
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
10
20
30
40
x

```
>         # codigo actuar:::hist.grouped.data
>
> var.group <- function (x, ...)
+ {
+   cj <- eval(expression(cj), envir = environment(x))
+   midpoints <- cj[-length(cj)] + diff(cj)/2
+   x <- as.matrix(x[-1])
+   drop(crossprod(x,midpoints^2)/colSums(x) - (crossprod(x, midpoints)/colSums(x))^2 )
+ }
>
>   # exemplo de trasformación unha táboa de frecuencias
>   xx <- runif(1000)
>   cj <- table(cut(xx,breaks=(0:10)/10))
>   x <- grouped.data( Group=(0:10)/10, Freq= as.vector(cj))
```

Moda (Mo): es el valor de la variable que más veces se repite, es decir, el más frecuente. La moda puede no ser única, puede haber una moda (variable unimodal), dos modas (bimodal), etc.

Cálculo de la moda:

- Distribuciones no agrupadas: valor de la variable de mayor frecuencia absoluta o relativa.
- Distribuciones agrupadas: el cálculo se realiza en dos estapas:
  - Detectar el *intervalo modal*: $[L_{i-1}, L_i)$ con mayor densidad de datos d_i.
  - Cálculo de la moda:

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i+1} + d_{i-1}} a_i$$

siendo L_{i-1}: el extremo inferior del intervalo modal,

d_{i-1} y d_{i+1}: las densidades de frecuencia de los intervalos anterior y posterior al modal respectivamente. (En caso de no existir intervalo anterior o posterior se consideran igual a 0),

a_i: la amplitud del intervalo modal.

Ejemplo 2.10. Calcula la moda en el ejercicio de la tabla agrupada del Ejercicio 2.1 Dado el vector de densidades ($d_1 = 0.5, d_2 = 4, d_3 = 12.5, d_4 = 0.5$), el intervalo (modal) con mayor densidad es $(28, 32]$. El valor de la moda es por lo tanto

$$Mo = 28 + \frac{0.5}{4 + 0.5} * 4 = 28 + \frac{2}{4.5} = 28 + 0.444 = 28.444$$

2.5 Medidas de posición no centrales: los Cuantiles

Se define *cuantil de orden p* con $0 < p < 1$ (x_p), como el valor que deja a lo sumo pN observaciones a su izquierda y $(1-p)N$ observaciones a su derecha. Destacamos en particular los cuantiles siguientes:

- Los *cuartiles* (Q_1, Q_2 y Q_3): dividen a la distribución en cuatro partes iguales, dentro de cada cual están incluidos el 25% de los valores de la distribución.

- Los *deciles* ($D_1, D_2, \ldots, D_9$): dividen a la distribución en diez partes iguales; dentro de cada una están incluidos el 10% de los valores.

- Los *percentiles* ($P_1, P_2, \ldots, P_{99}$): dividen a la distribución en cien partes iguales; dentro de cada una está incluido el 1% de los valores.

Cálculo del cuantil de orden p

1. Distribuciones no agrupadas: Se observa cuál es la primera frecuencia absoluta acumulada N_i que supera o iguala a pN distinguiéndose dos casos:
   - Si $N_i > pN$, entonces $x_p = x_i$.
   - Si $N_i = pN$, entonces $x_p = \dfrac{x_i + x_{i+1}}{2}$.
2. Distribuciones agrupadas: Primero se detecta el intervalo que contiene al cuantil: el primer $[L_{i-1}, L_i)$ con $N_i \geq pN$, luego se aplica la fórmula:

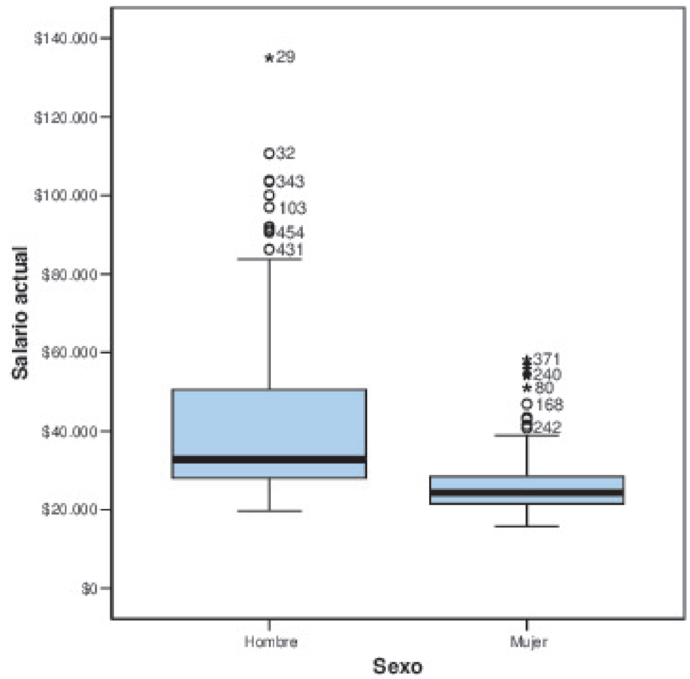
$$x_p = L_{i-1} + \frac{pN - N_{i-1}}{n_i} a_i = L_{i-1} + \frac{p - F_{i-1}}{f_i} a_i$$

Diagramas de caja

Este tipo de gráficos también se conocen como Box-Plots. Para construirlos es necesario el cálculo del primer y tercer cuartil Q_1 y Q_3. Se definen el extremo inferior y superior LI y LS respectivamente como:

$$LI = \max\left\{\min\{x_i\}, Q_1 - 3\left(\frac{Q_3 - Q_1}{2}\right)\right\} LS = \min\left\{\max\{x_i\}, Q_3 + 3\left(\frac{Q_3 - Q_1}{2}\right)\right\}$$

obteniéndose un diagrama de caja de la forma:



Ejercicio 2.5. Genera un diagrama de cajas de la variable Ancho del Pétalo en función del tipo de especie. Que conclusión en términos de los cuartiles de *Versicolor* podemos obtener?

Momentos Potenciales

Los momentos de una distribución son unos valores que la caracterizan, de tal modo que dos distribuciones son iguales si tienen todos sus momentos iguales, y son tanto más parecidas cuanto mayor sea el número de momentos iguales que tengan.

Momentos respecto al origen

El momento de orden r con respecto al origen se define como:

$$a_r = \frac{\sum_{i=1}^{k} x_i^r n_i}{N} = \sum_{i=1}^{k} x_i^r f_i$$

Los primeros momentos serán: $a_0 = 1$, $a_1 = \overline{X}$.

Momentos respecto a la media aritmética:

El momento de orden r con respecto a la media aritmética se define como:

$$m_r = \frac{\sum_{i=1}^{k}(x_i - \overline{x})^r n_i}{N} = \sum_{i=1}^{k}(x_i - \overline{x})^r f_i$$

Los primeros momentos serán: $m_0 = 1$, $m_1 = 0$, $m_2 = \dfrac{\sum_{i=1}^{k}(x_i - \overline{x})^2 n_i}{N} = s^2$.

Propiedad:

Utilizando el binomio de Newton, los momentos con respecto a la media se pueden expresar en función de los momentos con respecto al origen. En particular:

$$\begin{aligned} m_2 &= a_2 - a_1^2 \\ m_3 &= a_3 - 3a_2a_1 + 2a_1^2 \end{aligned}$$

Como ejercicio se propone expresar m_r los momentos respecto a la media aritmética de orden r en función de los momentos con respecto al origen.

2.6 Medidas de Dispersión

Son medidas que nos indican la desviación de los valores de la variable respecto de ciertas medidas de posición como la media aritmética o la mediana. A la mayor o menor separación de los valores respecto a otro, que se pretende sea su síntesis, se llama dispersión o variabilidad.

Medidas de Dispersión Absolutas:

Recorrido: es la diferencia entre el mayor valor y el menor valor de una distribución: $Re = x_k - x_1 = \max_{i=1 \ldots k} x_i - \min_{i=1 \ldots k} x_i$.

Recorrido Intercuartílico: es la diferencia existente entre el tercer cuartil y el primero: $IRQ = C_3 - C_1$.

Nos indica que en un intervalo de longitud IRQ están comprendidos el 50% de los valores centrales.

Varianza: De todas las medidas de dispersión absolutas respecto a la media aritmética, la varianza y su raíz cuadrada, la desviación típica son las más importantes.

$$s^2 = \frac{\sum_{i=1}^{k}(x_i - \overline{x})^2 n_i}{N}$$

Desviación típica o estándar: Es la raíz cuadrada de la varianza:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{k}(x_i - \overline{x})^2 n_i}{N}}$$

Propiedades de la varianza:

1. La varianza y la desviación típica **NUNCA** pueden ser negativas $s^2 \geq 0$, $s \geq 0$.

2. La varianza es la medida cuadrática de dispersión óptima, ya que:

$$\min_c \frac{1}{N} \sum_{i=1}^{k} (x_i - c)^2 n_i = \frac{1}{N} \sum_{i=1}^{k} (x_i - \overline{x})^2 n_i \text{ (Teorema de König)}$$

3. La varianza es igual al momento de segundo orden respecto al origen menos el de primer orden elevado al cuadrado.

$$s^2 = a_2 - a_1^2 = \left(\sum_{i=1}^{k} x_i^2 f_i \right) - \left(\sum_{i=1}^{k} x_i f_i \right)^2$$

4. Si en la distribución de frecuencias sumamos a todos los valores de la variable una constante, la varianza no varía (un cambio de origen no afecta a la varianza).

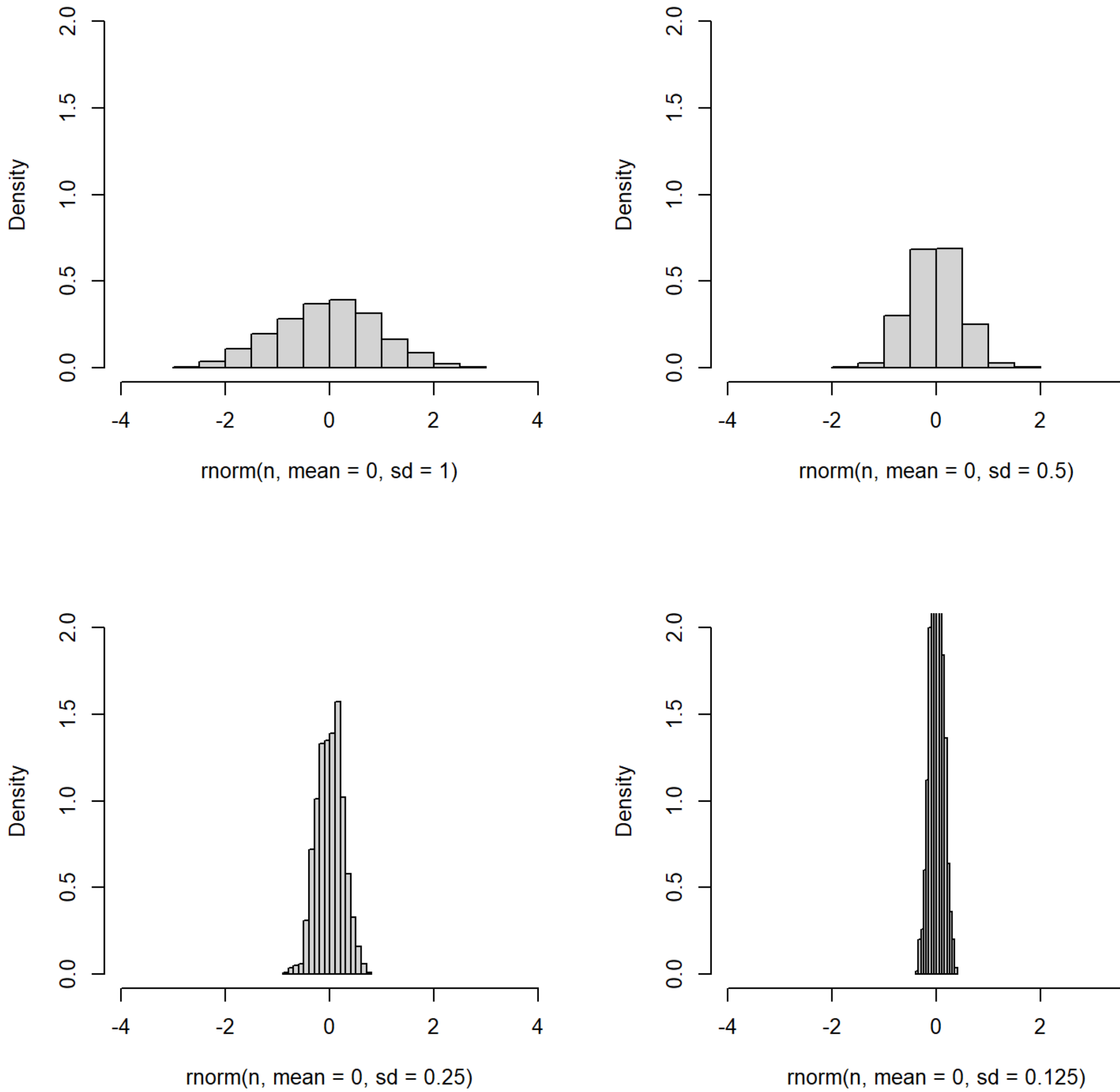
$$y_i = x_i + a \; ; \; s_Y^2 = s_X^2$$

5. Al multiplicar todos los valores de una distribución de frecuencias por una constante, la varianza queda multiplicada por el cuadrado de dicha constante.

$$y_i = bx_i \; ; \; s_Y^2 = b^2 s_X^2$$

Interpretación

```
> ### Código de R
> n <- 1000
> par( mfrow=c(2,2))
> hist( rnorm( n, mean=0, sd= 1), freq=FALSE , ylim=c(0,2), xlim=c(-4,4), main="")
> hist( rnorm( n, mean=0, sd= 0.5), freq=FALSE , ylim=c(0,2), xlim=c(-4,4), main="")
> hist( rnorm( n, mean=0, sd= 0.25), freq=FALSE , ylim=c(0,2), xlim=c(-4,4), main="")
> hist( rnorm( n, mean=0, sd= 0.125), freq=FALSE , ylim=c(0,2), xlim=c(-4,4), main="")
```


Density
0.0
0.5
1.0
1.5
2.0
-4
-2
0
2
4
rnorm(n, mean = 0, sd = 1)
rnorm(n, mean = 0, sd = 0.5)
rnorm(n, mean = 0, sd = 0.25)
rnorm(n, mean = 0, sd = 0.125)

```
> par( mfrow=c(1,1))
```

2.7 Medidas de Dispersión relativas

Si queremos comparar los promedios de dos distribuciones para saber cuál de los dos es más representativo debemos utilizar medidas adimensionales, es decir, que no vengan afectadas por las unidades de medida. Estas medidas de dispersión, llamadas relativas, siempre se concretan en forma de cociente. Las más utilizadas son:

Recorrido relativo: Se define como el cociente entre el recorrido y la media aritmética:

$$R_r = \frac{Re}{\overline{x}}$$

Coeficiente de variación de Pearson: se define como la relación por cociente entre la desviación típica y la media aritmética. Si $\overline{x} \neq 0$

$$v = \frac{s}{\overline{x}}$$

Cuanto mayor sea v, menor es la representatividad de la media aritmética.

Ejercicio 2.6. Dado el conjunto de datos *iris*, responde:

1. Agrupa la variable *Sepal.Length* en 8 intervalos de igual longitud. Usa el menú *Estadística Básica/Datos/Modificar el conjunto de datos activo/Segmentar variable numérica* .... Obtén su distribución de frecuencias completa.

2. Agrupa nuevamente pero en 8 intervalos con extremos 4 y 8. Usa el menú *Estadística Básica/Datos/Modificar el conjunto de datos activo/Recodificar* .... Obtén su distribución de frecuencias completa. Genera una nueva columna con la marca de clase

3. Compara la media da v. *Sepal.Length* agrupada y sin agrupar.

4. Comprueba la propiedad 5 de la media para la v. *Sepal.Length*, i.e., min(xi) <= media & media<= max(xi)

5. Recarga el conjunto de datos original y construye el histograma de la v. *Sepal.Length*.

6. Representa y compara en un gráfico la v. *Sepal.Length* en función del tipo de flor. Visualizando el gráfico, que se puede decir del 75% de las longitudes del Sepalo de la subespecie *setosa* con respecto a la subespecie *versicolor* y *virginica*. Y del 50% de la longitud del sepalo de la subespecie *virginica* con respecto a las otras.

7. Haz un resumen numérico de la v. *Sepal.Length* en función del tipo de subespecie.

8. Que proporción de valores de la v. *Sepal.Length* están dentro del intervalo $\bar{x} - 1.96 * sd_{Sepal.Length}, \bar{x} + 1.96 * sd_{Sepal.Length}$

9. Edita el conxunto de datos y pon el 1º valor de la v. *Sepal.Length* a 1000. Recalcula la media nuevamente. ¿Sigue siendo ese valor representativo de todas las longitudes del sepalo?.

10. ¿Qué valor representa ahora mejor la posición central?. Que ocurre si se pone el 2º valor= 1000?, y si se hace con el 3º valor también, e con el 4º, ... Cuantos valores se tienen que modificar para que se vea alterada esta nueva medida de posición central?.

11. Añade un valor para que la media de la variable *Petal.Length* sea 4.

```
> head(iris)
```

```
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5          1.4         0.2  setosa
2          4.9         3.0          1.4         0.2  setosa
3          4.7         3.2          1.3         0.2  setosa
4          4.6         3.1          1.5         0.2  setosa
5          5.0         3.6          1.4         0.2  setosa
6          5.4         3.9          1.7         0.4  setosa
```

```
> x <- mean(iris$Petal.Length) # A media dos (150+1) datos é: (sum_xi + x0)/151 =4, de aquí despe
>                         x0 <- 150*(4*151/150 - x); x0
```

```
[1] 40.3
```

```
>                         # Comprobamos que o valor x0 é o correcto
>                         mean( c(iris$Petal.Length,x0))
```

```
[1] 4
```

12. Modifica un valor para que la media sea igual a la mediana da v. *Petal.Length*

```
> x <- mean(iris$Petal.Length)
> y <- median(iris$Petal.Length)
> xx <- iris$Petal.Length
> xx[length(xx)] <- xx[length(xx)] + 150*(y-x)
> # Comprobamos que o valor x0 é o correcto
> mean(xx); y
```

```
[1] 4.35
```

```
[1] 4.35
```

Solución: *iris-resumen-numerico.html*

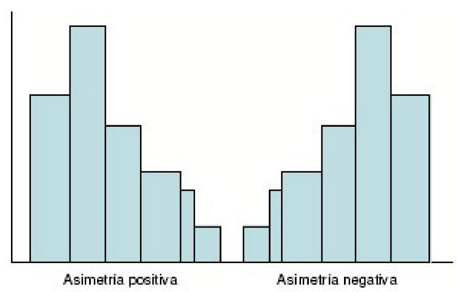
2.8 Medidas de Forma

Están relacionadas con la representación gráfica de la distribución. Pueden ser:

Medidas de simetría:

- Asimetría positiva: Si las frecuencias más altas se encuentran en el lado izquierdo de la media, mientras que en derecho hay frecuencias más pequeñas (cola).

- Asimetría negativa: Cuando la cola está en el lado izquierdo.



Coeficiente de asimetría de Fisher: $g_1 = \dfrac{m_3}{s^3}$

Para cada coeficiente, existen tres posibilidades: - $g_1 > 0$: La asimetría es positiva o por la derecha.

- $g_1 = 0$: La distribución es simétrica.

- $g_1 < 0$: La asimetría es negativa o por la izquierda.

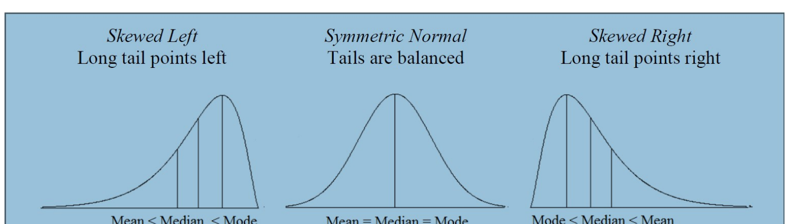


Figure 1. Sketches showing general position of mean, median, and mode in a population.

Figura: Extraído de: Doane, D.P., Seward, L.E. (2011). Measuring Skewness: A Forgotten Statistic?. *Journal of Statistics Education*, Volume **19**, Number 2.

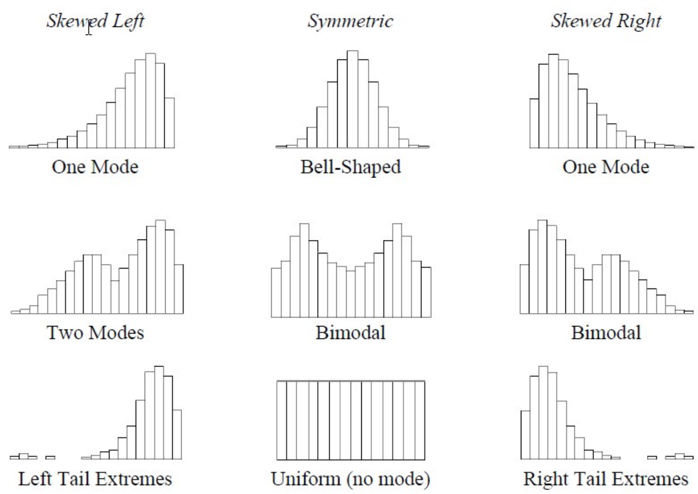
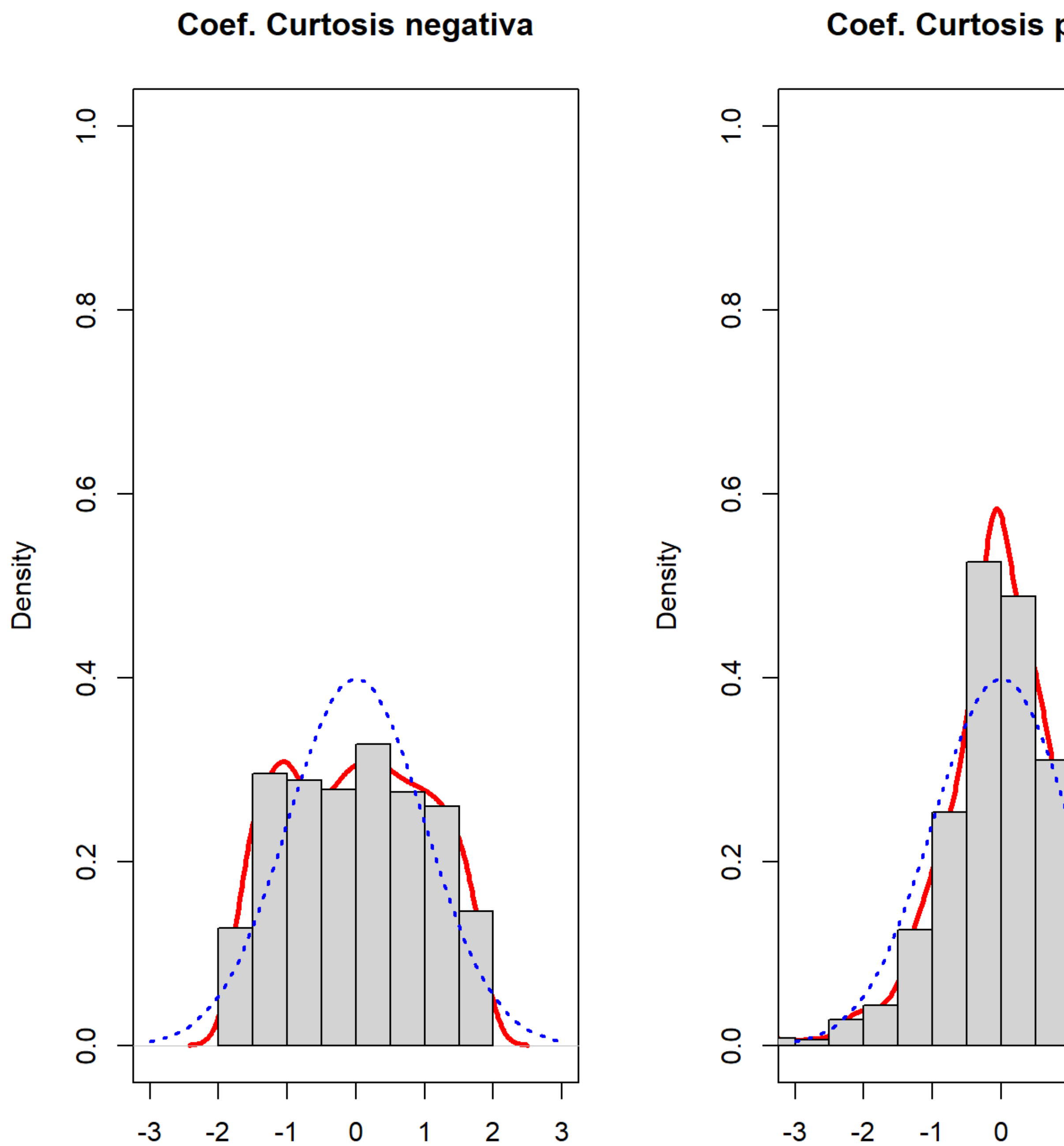


Figure 2. Illustrative prototype histograms.

Medidas de apuntamiento o curtosis

- Coeficiente de curtosis de Fisher: $g_2 = \dfrac{m_4}{s^4} - 3$
  - $g_2 > 0$: Distribución leptocúrtica: más apuntamiento que la distribución normal.
  - $g_2 = 0$: Distribución mesocúrtica: apuntamiento similar a la distribución normal.
  - $g_2 < 0$: Distribución platicúrtica: menos apuntamiento que la distribución normal.


Coef. Curtosis negativa
Density
0.0
0.2
0.4
0.6
0.8
1.0
-3
-2
-1
0
1
2
3
Coef. Curtosis p
Density
0.0
0.2
0.4
0.6
0.8
1.0
-3
-2
-1
0

Tipificación

Una variable estadística se dice tipificada o estandarizada si su media es cero y su varianza o su desviación típica es uno. Dada una variable X con media μ, y varianza σ^2, la nueva variable $Z = \dfrac{X-\mu}{\sigma}$, es su tipificada.

Ejercicio 2.7. Tipifica la variable *Sepal.Length* del conjunto de datos *iris* y comprueba que efectivamente la v. tipificada tiene media 0 y varianza 1. Menú *Estadística Básica/Datos/Modificar variables del conjunto de datos activo/Tipificar*.

```
> data(iris)
> iris <- local({
+   .Z <- scale(iris[,c("Sepal.Length")])
+   within(iris, {
+     Z.Sepal.Length <- .Z[,1]
+   })
+ })
>
> RcmdrMisc::numSummary(iris[,"Z.Sepal.Length"], statistics=c("mean", "sd", "IQR", "quantiles"),
+   quantiles=c(0,0.25,0.5,0.75,1))
```

```
          mean sd      IQR       0%        25%         50%      75%     100%
 -4.480675e-16  1 1.569923 -1.86378 -0.8976739 -0.05233076 0.672249 2.483699
   n
 150
```