Estadística Descriptiva

Examinaremos los datos en forma descriptiva para:

- Organizar la información
- Sintetizar la información
- Ver sus características más relevantes
- Presentar la información

de datos provenientes de una o varias muestras de toda la población.

El objetivo con frecuencia es obtener conclusiones sobre toda la población a partir de una muestra

→ *Métodos de Estadística de Inferencia:*

-Estimación Puntual
-Estimación por Intervalos
-Tests de Hipótesis

Factores necesarios para un buen análisis estadístico:

- Diseño del Experimento o Investigación
- Calidad de los Datos

Población: todos los individuos que poseen la(s) característica(s) de interés.

Muestra: subconjunto de la población

Población ----------> Muestra
<----------
Inferencia

Las siguientes son mediciones de la proporción de la masa de la Tierra con respecto a la Luna

Mariner II	81.3001
Mariner IV	81.3015
Mariner V	81.3006
Mariner VI	81.3011
Mariner VII	81.2997
Pioneer VI	81.3005
Pioneer VII	81.3021

En *Probabilidad*, por ejemplo, podríamos suponer que las posibles mediciones se distribuyen alrededor del verdadero valor81.3035 y nos preguntaríamos

¿Cuál es la probabilidad de que las 7 mediciones sean menores que el verdadero valor?

En *Estadística,* a partir de los 7 observaciones nos preguntaríamos:

¿Muestran los datos que el verdadero valor del cociente es 81.3035?
¿Cuán confiable es decir que el verdadero valor está en (81.2998, 81.3018?)

Las técnicas del análisis exploratorio nos ayudan a organizar la información que nos dan los datos, de manera de detectar algún patrón de comportamiento así como también apartamientos importantes al modelo subyacente. Nos guían a la estructura subyacente en los datos de manera rápida y simple.

Gráficos de Tallo y Hoja

Nos dan una primera aproximación rápida a la distribución de los datos sin perder de vista las observaciones.

1. Separamos a cada observación en dos partes: **tallo** y **hoja**

2. Listamos en forma vertical y creciente los tallos y agregamos las hojas a la derecha del tallo correspondiente.

Ejemplo: La siguiente tabla muestra los datos de la fuerza compresión de 45 muestras de aleación de Aluminio-Litio.

96	93	88	117	127	95	113	96
108	94	148	156	139	142	94	107
125	155	155	103	112	127	117	120
112	135	132	111	125	104	106	139
134	119	97	89	118	136	125	143
120	103	113	124	138			

Ordenamos los datos

88	89	93	94	94	95	96	96
97	103	103	104	106	107	108	111
112	112	113	113	117	117	118	119
120	120	124	125	125	125	127	127
132	134	135	136	138	139	139	142
143	148	155	155	156			

1. Elegimos un número de dígitos: 2 en este caso
2. Separamos los dígitos de los restantes, nos quedan 8 tallos de 8 a 15.
3. Obtenemos las hojas

```
 8 | 8 9
 9 | 3 4 4 5 6 6 7
10 | 3 3 4 6 7 8
11| 1 2 2 3 3 7 7 8 9
12| 0 0 4 5 5 5 7 7
13| 2 4 5 6 8 9 9
14 | 2 3 8
15 | 5 5 6
```

¿Qué podemos ver en este diagrama?

- Forma de la distribución: simetría, asimetría a derecha, asimetría a izquierda
- Posición del centro de la distribución y concentración de los datos
- Desviaciones marcadas respecto al comportamiento general: outliers

Ejemplo:

Los siguientes datos corresponden a tiempos de falla de de cables Kevlar 49/epoxy sometidos a una presión del 90%:

TIEMPOS DE FALLA

0.01 0.01 0.02 0.02 0.02 0.03 0.03 0.04 0.05 0.06 0.07 0.07 0.08 0.09 0.09 0.10
0.10 0.11 0.11 0.12 0.13 0.18 0.19 0.20 0.23 0.80 0.80 0.83 0.85 0.90 0.92 0.95
0.99 1.00 1.01 1.02 1.03 1.05 1.10 1.10 1.11 1.15 1.18 1.20 1.29 1.31 1.33 1.34
1.40 1.43 1.45 1.50 1.51 1.52 1.53 1.54 1.54 1.55 1.58 1.60 1.63 1.64 1.80 1.80
1.81 2.02 2.05 2.14 2.17 2.33 3.03 3.03 3.24 4.20 4.69 7.89

El esquema de tallo y hoja resulta:

```
STEM AND LEAF PLOT OF PER90

   LEAF DIGIT UNIT = 0.1           MINIMUM  0.0100
   7  8  REPRESENTS 7.8            MEDIAN   1.0750
                                   MAXIMUM  7.8900
         STEM  LEAVES
     25    0  00000000000000001111111122
     33    0  88889999
    (18)   1  0000011111122333444
     25    1  55555555666888
     11    2  00113
      6    2
      6    3  002
      3    3
      3    4  2
      2    4  6
      1    5
      1    5
      1    6
      1    6
      1    7
      1    7  8
```

En este caso cada tallo ha sido dividido en 2 líneas.

* 0, 1, 2, 3, 4
* 5, 6, 7, 8, 9

Se observa asimetría a derecha y un valor alejado del resto: 7.8

¿Qué significan los números en la columna de la izquierda?
Es la Profundidad

A cada dato le podemos asignar un valor de ranking o rango contando desde cada extremo de la muestra ordenada. La **profundidad** es el menor de los dos valores.

En el *stem and leaf plot* el número en la columna de la izquierda es la mayor profundidad de la línea, excepto en aquella en la que el número está entre paréntesis, pues en ese caso el número que figura es la cantidad de hojas que hay en dicha línea.

Veamos otro ejemplo

Ejemplo: Concentración de Inmunoglobulina en 298 niños sanos entre 6 meses y 6 años de edad.

Igm	**n°. de niños**
0.1	3
0.2	7
0.3	19
0.4	27
0.5	32
0.6	35
0.7	38
0.8	38
0.9	22
1.0	16
1.1	16
1.2	6
1.3	7
1.4	9
1.5	6
1.6	2
1.7	3
1.8	3
2.0	3
2.1	2
2.2	1
2.5	1
2.7	1
4.5	1

Veamos el esquema de tallo y hoja que construye el SX.

```
 LEAF DIGIT UNIT = 0.1                MINIMUM  0.1000
0  1  REPRESENTS 0.1                 MEDIAN   0.7000
                                     MAXIMUM  4.5000
       STEM  LEAVES
     3    0  111
    29    0  22222223333333333333333333
    88    0  44444444444444444444444444555555555555555555555555*
   (73)   0  66666666666666666666666666666666677777777777777777*
   137    0  88888888888888888888888888888888888899999999999999*
    77    1  00000000000000001111111111111111
    45    1  2222223333333
    32    1  444444444555555
    17    1  66777
    12    1  888
     9    2  00011
     4    2  2
     3    2  5
     2    2  7
     1    2
     1    3
     1    3
     1    3
     1    3
     1    3
     1    4
     1    4
     1    4  5
298 CASES INCLUDED   0 MISSING CASES
```

En este caso cada tallo ha sido dividido en 5 líneas.

*	0 y 1
t	2 y 3
f	4 y 5
s	6 y 7
•	8 y 9

Hay reglas heurísticas para elegir el número de tallos. En general estos van entre 8 y 20.

Cuando el volumen de datos es muy grande conviene usar otro tipo de gráficos que también son de fácil interpretación .

Histogramas

- Dividimos el rango donde viven los datos **n** en **intervalos o clases**, que no se superpongan. Las clases deben ser **excluyentes** y **exhaustivas**.

- Contamos la cantidad de datos en cada intervalo o clase, es decir la **frecuencia**. También podemos usar para cada intervalo la **frecuencia relativa**

$$fr_i = \frac{f_i}{n}$$

- Graficamos el histograma en un par de ejes coordenados representando en las abscisas los intervalos y sobre cada uno de ellos un rectángulo cuya área es proporcional a la frecuencia relativa de dicho intervalo.

Obs.:

- No existen criterios óptimos para elegir la cantidad de intervalos. En general, entre 8 y 15 intervalos deberían ser suficientes. Muchos o muy pocos intervalos puede ser poco informativo.
  Se busca un equilibrio entre un histograma muy irregular y uno demasiado suavizado.

- No es necesario que todos los intervalos tengan la misma longitud, pero es recomendable que así sea. Esto facilita la lectura.

- El histograma representa la frecuencia o la frecuencia relativa a través del **área** y no a través de la altura.

- Es recomendable tomar

$$\textbf{Altura del rectángulo} = \frac{\textbf{frecuencia relativa}}{\textbf{Long. del intervalo}}$$

De esta manera el área es 1 y dos histogramas son fácilmente comparables independientemente de la cantidad de observaciones en las que se basa cada uno.

Ejemplo: Porcentajes de octanos para mezclas de naftas.

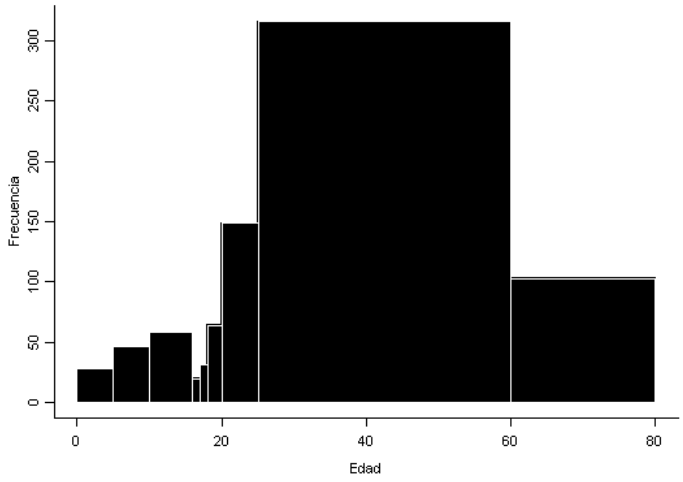
85.3	87.5	87.8	88.5	89.9	90.4	91.8	92.7
86.7	87.8	88.2	88.6	90.3	91.0	91.8	93.2
88.3	88.3	89.0	89.2	90.4	91.0	92.3	93.3
89.9	90.1	90.1	90.8	90.9	91.1	92.7	93.4
91.2	91.5	92.6	92.7	93.3	94.2	94.7	94.2
95.6	96.1						

Clase	Frecuencia f_i	Frecuencia relativa fr_i
(85, 87]	2	0.048
(87, 89]	9	0.214
(89, 91]	12	0.286
(91,93]	10	0.238
(93,95]	7	0.167
(95,97]	2	0.048
Total	**42**	**1**

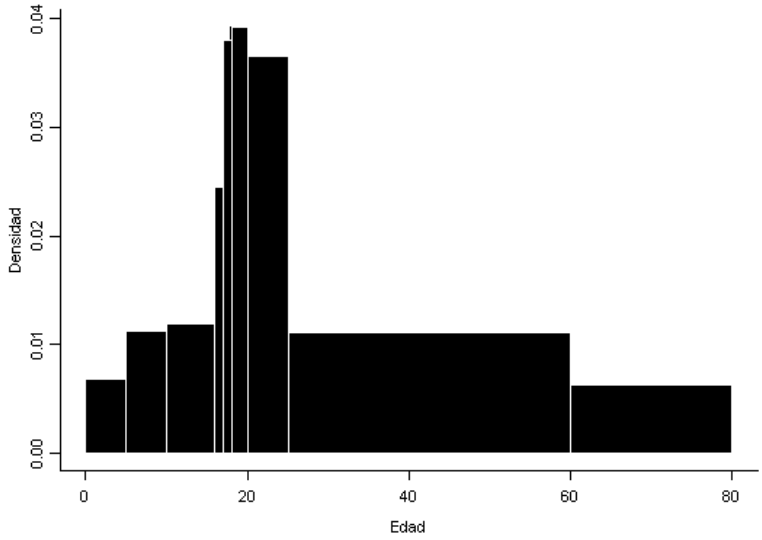
Ejemplo: Supongamos que la información que tenemos viene dada por la siguiente tabla:

Edad	Frecuencia	
0-4	28	
5-9	46	
10-15	58	Víctimas de accidentes automovilísticos en Londres en 1985
16	20	
17	31	
18-19	64	
20-24	149	
25-59	316	
60-80	103	
Total	**815**	

¿Qué observaríamos si graficásemos las frecuencias?



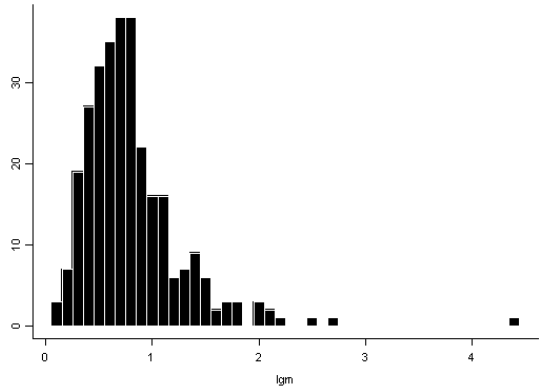
Histograma MAL HECHO !!!



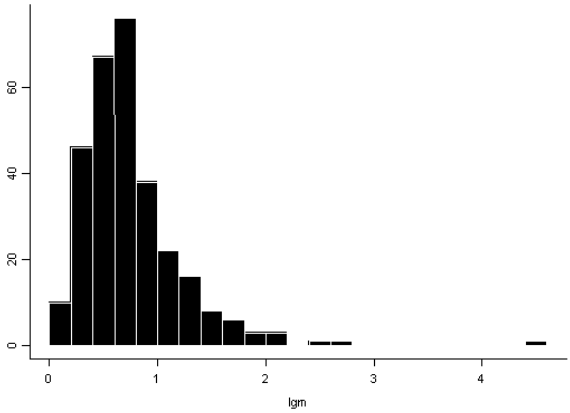
La forma correcta de graficarlo es:

La escala vertical es la densidad, es decir la frecuencia relativa dividida la longitud de la clase correspondiente. Si tuviéramos individuos accidentados parados en cada grupo etáreo, la altura del histograma representaría el aglutinamiento en cada clase: **hay partes del eje de abscisas que están más densamente pobladas que otras.**

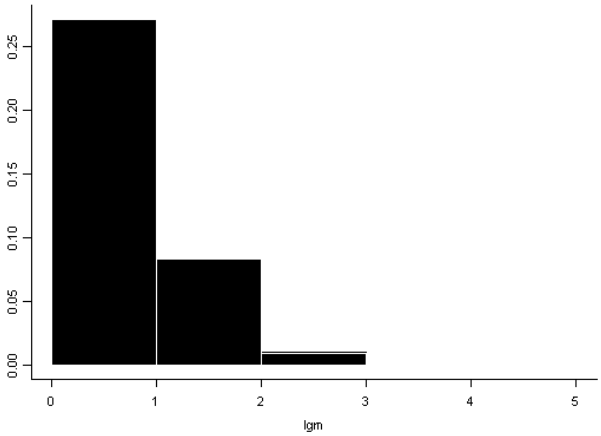
Ejemplo: Concentración de Img



Longitud de Clase= 0.1 g/l



Longitud de Clase= 0.2 g/l



Longitud de Clase=1g/l

Asimetría

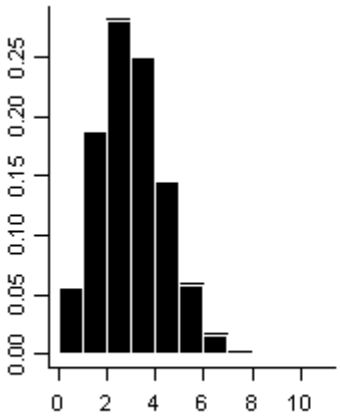
Un conjunto de datos que no se distribuye simétricamente, se llama **asimétrico**.

La asimetría puede verse en el esquema de Tallo y Hoja o en el Histograma.

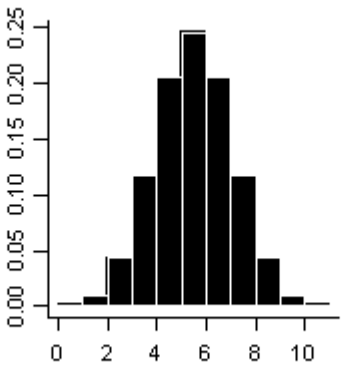
También se puede apreciar a través de la posición relativa entre media y mediana.

En un boxplot lo haremos a través de la posición relativa entre la mediana y los cuartiles.

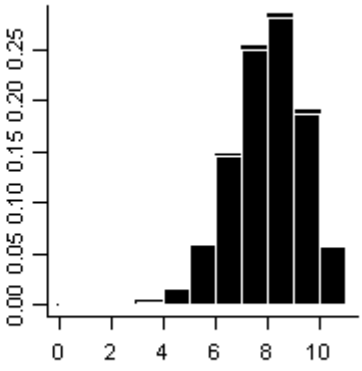
Un histograma tendería a tener la siguientes formas según cada caso:



Asimétrica a derecha



Simétrica



Asimétrica a Izquierda

Medidas de Resumen

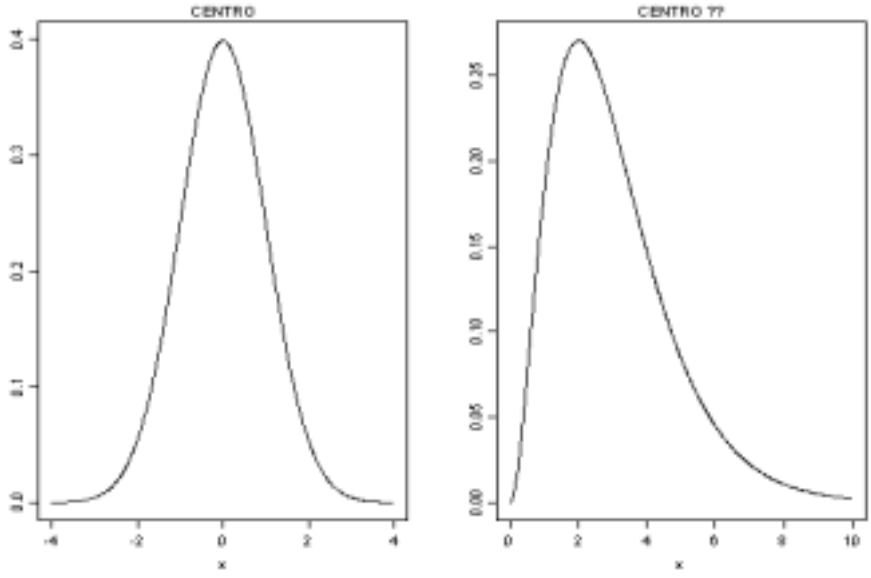
Resumiremos la información de los datos mediante medidas de fácil interpretación que reflejen sus características más relevantes.

La mediada a elegir dependerá de cada problema.

Medidas de Posición o Centrado

¿Cuál es el valor central o que mejor representa a los datos?
Buscamos un valor típico que represente a los datos.

Si la distribución es simétrica diferentes medidas darán resultados similares. Si es asimétrica no existe un centro evidente y diferentes criterios para resumir los datos pueden diferir considerablemente, en tanto tratan de captar diferentes aspectos de los mismos.



Promedio o Media Muestral

- Sumamos todas las observaciones y dividimos por el número total datos.

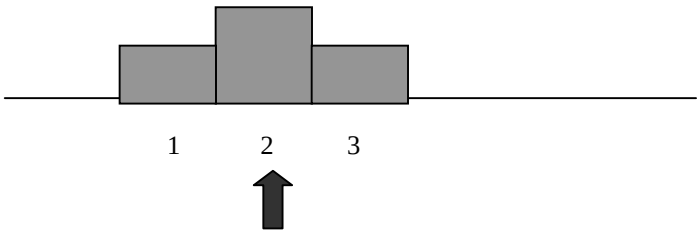
$$\bar{x} = \frac{x_1 + x_2 + \ldots\ldots + x_n}{n}$$ Promedio o Media Muestral

Ejemplo: Fuerza de compresión de muestras de Aleación de Aluminio-Litio

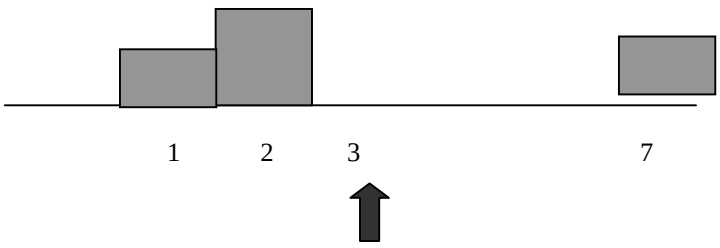
$$\bar{x} = \frac{\sum_{i=1}^{45} x_i}{45} = \frac{5350}{45} = 118.89$$

Es el punto de equilibrio del conjunto de datos.

X´s: 1, 2, 2, 3



X´s: 1, 2, 2, 7



Es una medida muy sensible a la presencia de datos anómalos (outliers).

Mediana Muestral

Es una medida del centro de los datos en tanto divide a la muestra ordenada en dos partes de igual tamaño. Deja la mitad de los datos a cada lado.

Sean los estadísticos de orden muestrales:

$$x_{(1)} \leq x_{(2)} \leq \ldots\ldots \leq x_{(n)}$$

definamos como mediana

$$\tilde{x} = \begin{cases} x_{(k=1)} & \text{si} \quad n = 2k+1 \\ \dfrac{x_{(k)} + x_{(k+1)}}{2} & \text{si} \quad n = 2k \end{cases}$$

Si la distribución es simétrica la mediana y la media identifican al mismo punto.

La mediana es resistente a la presencia de datos atípicos.

Puede ser útil cuando algunos datos han sido censurados.

Si tenemos:

X´s: 1,2,2,3 $\quad \bar{x} = 2 \quad \tilde{x} = 2$

X´s: 1,2,2,7 $\quad \bar{x} = 3 \quad \tilde{x} = 2$

¿Qué pasa si tenemos un 70 en lugar de 7?

Si tenemos una muestra de salarios de una población dada, ¿sería más adecuado tomar la media o la mediana muestral para representarlos?

Medias α-Podadas

Es un promedio calculado sobre los datos una vez que se han eliminado **α %** de los datos más pequeños y **un α %** de los datos más grandes. Formalmente podemos definirla como:

$$\overline{x}_{\alpha} = \frac{x_{([n\alpha]+1)} + \ldots + x_{(n-[n\alpha])}}{n - 2[n\alpha]}$$

La mediana puede ser vista como una 50% media podada.

Es una medida intermedia entre la media y la mediana

Es más resistente a datos atípicos que la media.

Veamos un ejemplo en el que calculamos las tres medidas

Los datos en la siguiente tabla corresponden al número de pulsaciones por minuto en pacientes con asma durante un espasmo:

Ordenamos los datos:

40 120 120 125 136 150 150 150 150 167

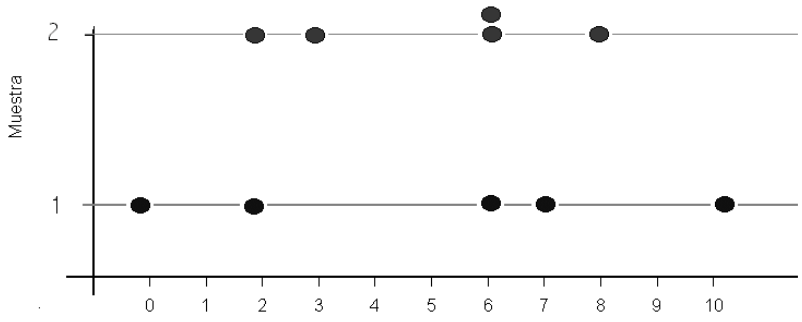
$\overline{x} = 130.8$ $\quad$ $\tilde{x} = 143$ $\quad$ $\overline{x}_{10} = 137.625$

Medidas de Dispersión o Variabilidad

¿Cuán dispersos están los datos? ¿Cuán cercanos son los datos al valor típico?

Supongamos que tenemos datos x_1 ,x_2 ,.....,x_n.

Veamos un ejemplo:



X´s:

0 2 6 7 10

Y´s:

2 3 6 6 8

$\overline{X} = \overline{Y} = 5$

$\tilde{X} = \tilde{Y} = 6$

¿Cómo medir la diferencia que se observa entre ambas muestras?

Rango Muestral

Es la diferencia entre el valor más grande y el pequeño de los datos:

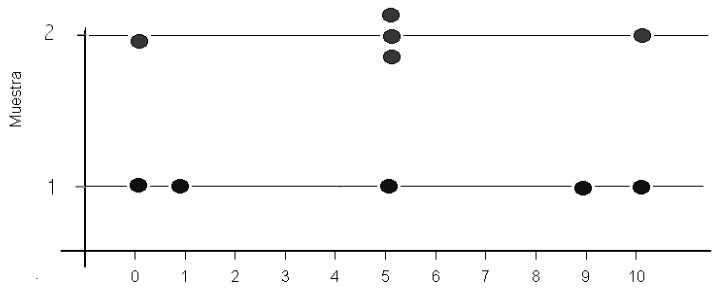
$$R_X = \text{rango} = \text{máx}(x_i) - \text{mín}(x_i).$$

Ejemplo: en nuestros conjuntos de datos:

$R_X = 10 \qquad R_Y = 6$

- Esta medida es muy sensible a la presencia de outliers.

Veamos otro ejemplo:



X´s:

0 1 5 9 10

Y´s:

0 5 5 5 10

$\bar{X} = \bar{Y}$

$\tilde{X} = \tilde{Y}$

$R_X = R_Y$

Varianza Muestral

Mide la variabilidad de los datos alrededor de la media muestral.

Varianza muestral $= S^2 = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

desvío estándar muestral $= S = \sqrt{S^2}$

Ejemplo: en los dos ejemplos anteriores obtenemos

$$S^2_x = 20.5 \qquad S_x = 4.258$$

$$S^2_y = 12.5 \qquad S_y = 3.536$$

- El desvío estándar tiene las mismas unidades que los datos, mientras que la varianza no.

- Al basarse en promedios, es sensible a la presencia de datos atípicos. Por ejemplo, si en la muestra de los Y´s cambiamos el 10 por un 15 obtenemos $\mathbf{S^2_y = 30}$ **y** $\mathbf{S_y = 5.477}$, mientras que si lo cambiamos por un 20 obtenemos $\mathbf{S^2_y = 57.5}$ **y** $\mathbf{S_y = 7.583}$.

Coeficiente de Variación:

Es una medida que relaciona el desvío standard con la media de una muestra:

$$CV = \frac{S}{\overline{X}}$$

Es una medida que está en desuso, ya que no tiene propiedades estadísticas muy interesantes, sin embargo no depende de las unidades y si lo multiplicamos por 100 nos da una idea de la variabilidad relativa.

Distancia Intercuartil

Es una medida más resistente que el desvío estándar.

Comenzaremos por definir los **percentiles.**

El percentil α % de la distribución de los datos es el valor por debajo del cual se encuentran el α % de los datos en la muestra ordenada.

Para calcularlo:

- Ordenamos la muestra de menor a mayor
- Buscamos el dato que ocupa la posición $\frac{\alpha(n+1)}{100}$ (si este número no es entero se promedian los dos adyacentes o se interpolan los dos adyacentes)

Ejemplo: Tenemos 19 datos que ordenados son

1 1 2 2 3 4 4 5 5 6 7 7 8 8 9 9 10 10 11

Percentil	**Posición**	**Valor**	
10%	10 (19+1)/100=**2**	1	
25%	25 (19+1)/100=**5**	3	**Cuartil Inferior**
50%	50 (19+1)/100=**10**	6	**Mediana**
75%	75 (19+1)/100=**15**	9	**Cuartil Superior**
95%	95 (19+1)/100=**19**	11	

El percentil 50% coincide con la mediana

Llamamos cuartil inferior al percentil 25% y cuartil inferior al percentil 75%.
Los cuartiles y la mediana dividen a la muestra ordenada en cuatro partes igualmente pobladas. Entre los cuartiles se hallan aproximadamente el 50% central de los datos y el rango de estos es:

d_I =distancia intercuartil
= cuartil superior-cuartil inferior

Observación: Si en ejemplo cambiáramos el último dato por 110, la distancia intercuartil no cambiaría, mientras que el desvío pasaría de 3.2 a 24.13!!!!

Desvío Absoluto Mediano (Desviación absoluta respecto de la Mediana)

$$MAD = \text{med}\,|x_i - \text{med}(x_i)|$$

Es una versión robusta del desvío estándar basada en la mediana.

Observación: Si deseamos comparar la distancia intercuartil y la MAD con el desvío standard es conveniente dividirlas por constantes adecuadas. En ese caso se compara a S con

$$\frac{MAD}{0.675} \qquad \frac{d_I}{1.35}$$

Ejemplo: En la siguiente tabla se muestran las mediciones de FEV1 (volumen expiratorio forzado en 1 segundo) en 13 pacientes adolescentes que padecen asma.

Paciente	FEV1	Paciente	FEV1
1	2.30	8	2.25
2	2.15	9	2.68
3	3.50	10	3.00
4	2.60	11	4.02
5	2.75	12	2.85
6	2.82	13	3.38
7	4.05		

DESCRIPTIVE STATISTICS

	FEV1
N	13
MEAN	2.9500
SD	0.6231
VARIANCE	0.3883
C.V.	21.123
MINIMUM	2.1500
1ST QUARTI	2.4500
MEDIAN	2.8200
3RD QUARTI	3.4400
MAXIMUM	4.0500
MAD	0.5200

$$\frac{MAD}{0.675} = 0.77$$

$$\frac{d_I}{1.35} = \frac{3.44 - 2.45}{1.35} = \frac{0.99}{1.35} = 0.733$$

5 Números de Resumen

Los 5 números de resumen de la distribución de un conjunto de datos consisten en el **mínimo**, el **cuartil inferior**, la **mediana**, el **cuartil superior** y el **máximo**.

Box-Plots

Con la medidas anteriores podemos construir un gráfico de fácil realización y lectura.
¿Cómo lo hacemos?

1. Representamos una escala vertical u horizontal

2. Dibujamos una caja cuyos extremos son los cuartiles y dentro de ella un segmento que corresponde al a mediana.

3. A partir de cada extremo dibujamos un segmento hasta el dato más alejado que está a lo sumo 1.5 $\mathbf{d_I}$ del extremo de la caja. Estos segmentos se llaman bigotes.

4. Marcamos con ***** a aquellos datos que están entre 1.5 $\mathbf{d_I}$ y 3 $\mathbf{d_I}$ de cada extremo y con **o** a aquellos que están a más de 3 $\mathbf{d_I}$ de cada extremo.

Observación: Muchos paquetes estadísticos realizan el boxplot usando la distancia intercuartil y otros usan la distancia entre cuartos. Como estas medidas son muy próximas, en general los resultados son análogos. Lo importante es que entre los cuartos o entre los cuartiles yace aproximadamente el 50% central de los datos.

Ejemplo:

Si tenemos los siguientes datos ya ordenados:

10	25	50	91	92
108	109	113	114	115
120	126	132	133	141
146	151			

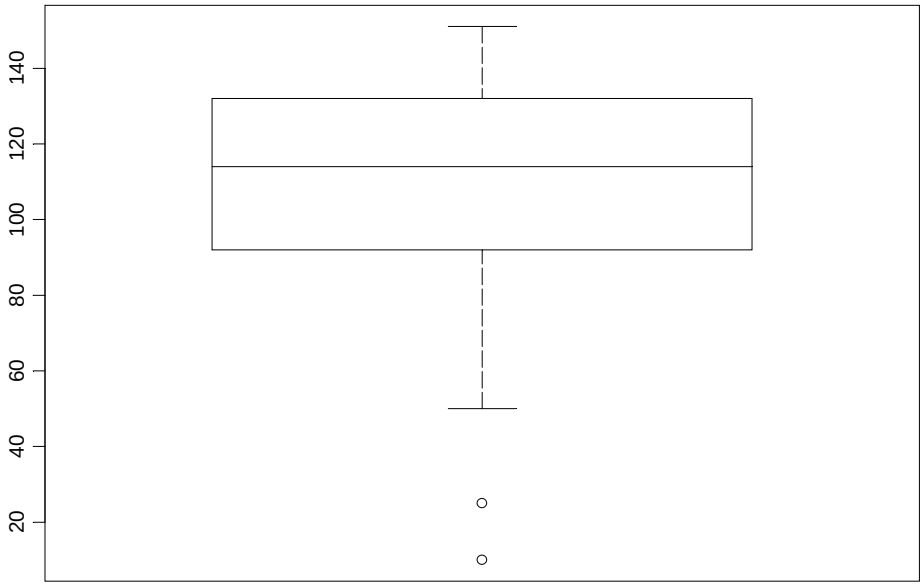
Cuartil inferior= 92 Cuartil superior= 132

$\mathbf{d_I}$ = 40 1.5 $\mathbf{d_I}$ **= 60** 3 $\mathbf{d_I}$ = 120

```
STEM AND LEAF PLOT OF X

   LEAF DIGIT UNIT = 10                MINIMUM  10.000
   0  1  REPRESENTS 10.                MEDIAN   114.00
                                       MAXIMUM  151.00
          STEM  LEAVES
        1    0  1
        2    0  2
        3    0  5
        3    0
        5    0  99
       (5)   1  00111
        7    1  2233
        3    1  445

17 CASES INCLUDED   0 MISSING CASES
```



A partir de un box-plot podemos podemos apreciar los siguientes aspectos de la distribución de un conjunto de datos:

- posición
- dispersión
- asimetría
- longitud de las colas
- puntos anómalos o outliers.

Los box-plots son especialmente útiles para comparar varios conjuntos de datos, pues nos dan una rápida impresión visual de sus características.

Outliers

Los métodos que hemos visto nos permiten identificar puntos atípicos, que pueden aparecer en una o más variables. Su detección es importante pues pueden determinar o

influenciar fuertemente los resutados de un análisis estadístico clásico, pues muchas de las técnicas habitualmente usadas son muy sensibles a la presencia de datos atípicos.

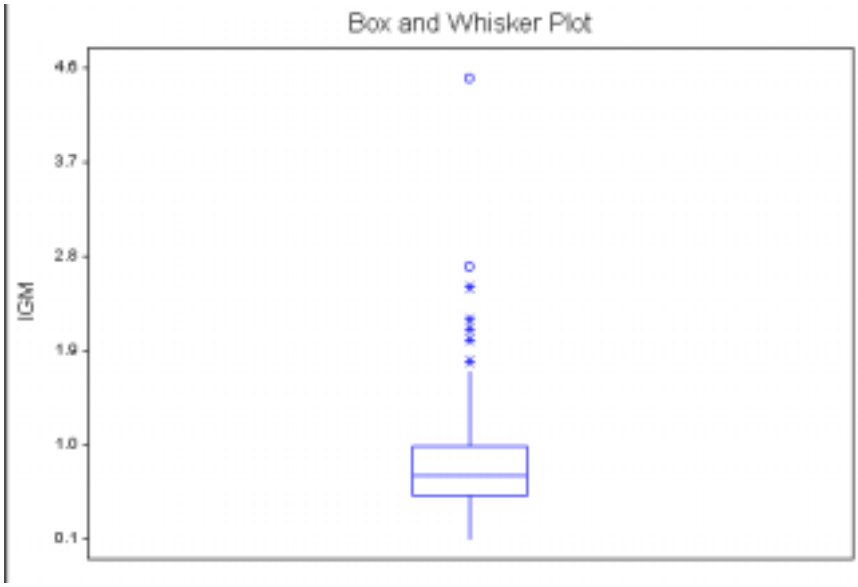
Los outliers deben ser cuidadosamente inspeccionados, si no hay evidencia de error y su valor es posible no deberían ser eliminados. Sin embargo, si el individuo tiene algo particular, como una enfermedad, su inclusión debería ser reconsiderada.

Podemos tener una idea de cuán influyentes son los datos.

Asimismo,la presencia de outliers puede indicar que la escala elegida no es la más adecuada.

Otros ejmplos:

- Igm



- Comparación de la población de la 10 ciudades más grandes en 16 paises

QQ-plot

El qq-plot es un gráfico que nos sirve para evaluar la cercanía a la distribución normal.

Para realizarlo se consideran los estadísticos de orden

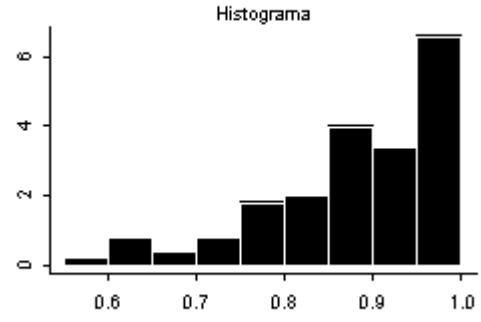
$$x_{(1)} \leq x_{(2)} \leq \ldots\ldots \leq x_{(n)}$$

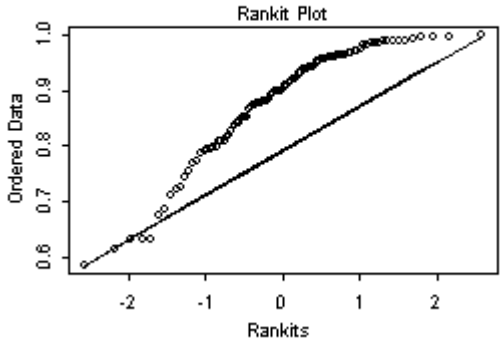
que se grafican versus el percentil $\frac{i-1/3}{n+1/3}$ de la normal, es decir $\phi^{-1}(\frac{i-1/3}{n+1/3})$.

Si los datos provienen de una distribución normal esperamos que el gráfico sea parecido a una recta.

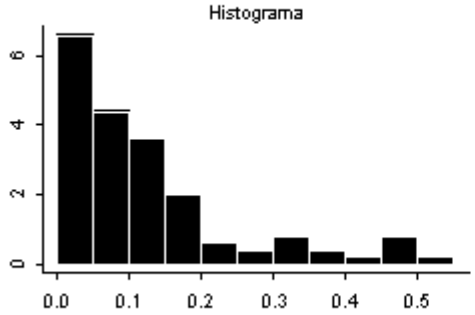
El alejamiento de la normalidad se ve reflejado por la forma del gráfico.

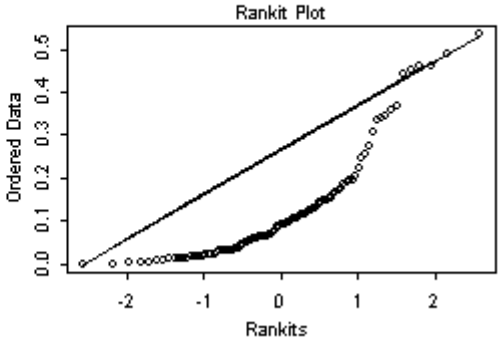
Asimetrica a Izquierda



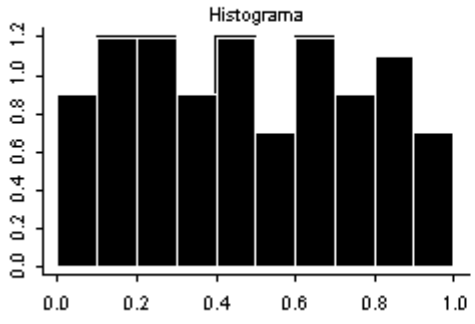


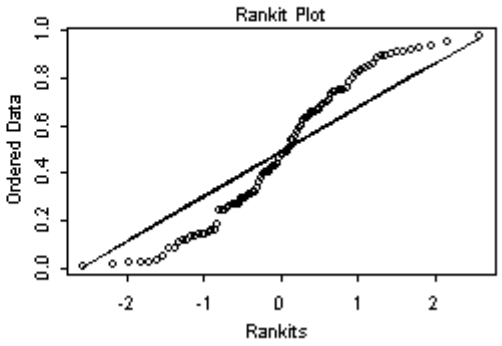
Asimetrica a Derecha



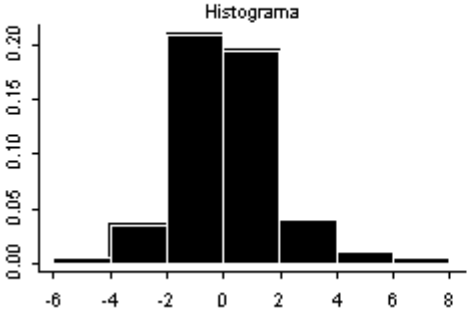


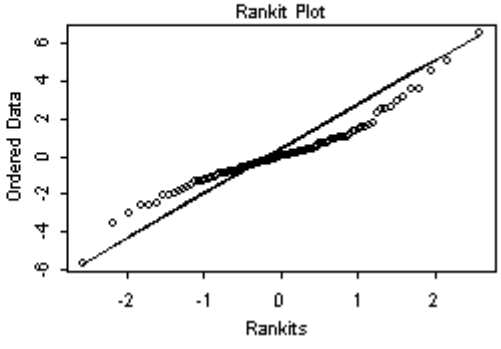
Simétrica con Colas Livianas





Simétrica con Colas Pesadas





Tiempos de CPU (en segundos)de 25 trabajos enviados a un server tomados al azar

				CPU
1.17	1.23	0.15	0.19	0.92
1.61	3.76	2.41	0.82	0.75
1.16	1.94	0.71	0.47	2.59
1.38	0.96	0.02	2.16	3.07
3.53	4.75	1.59	2.01	1.40

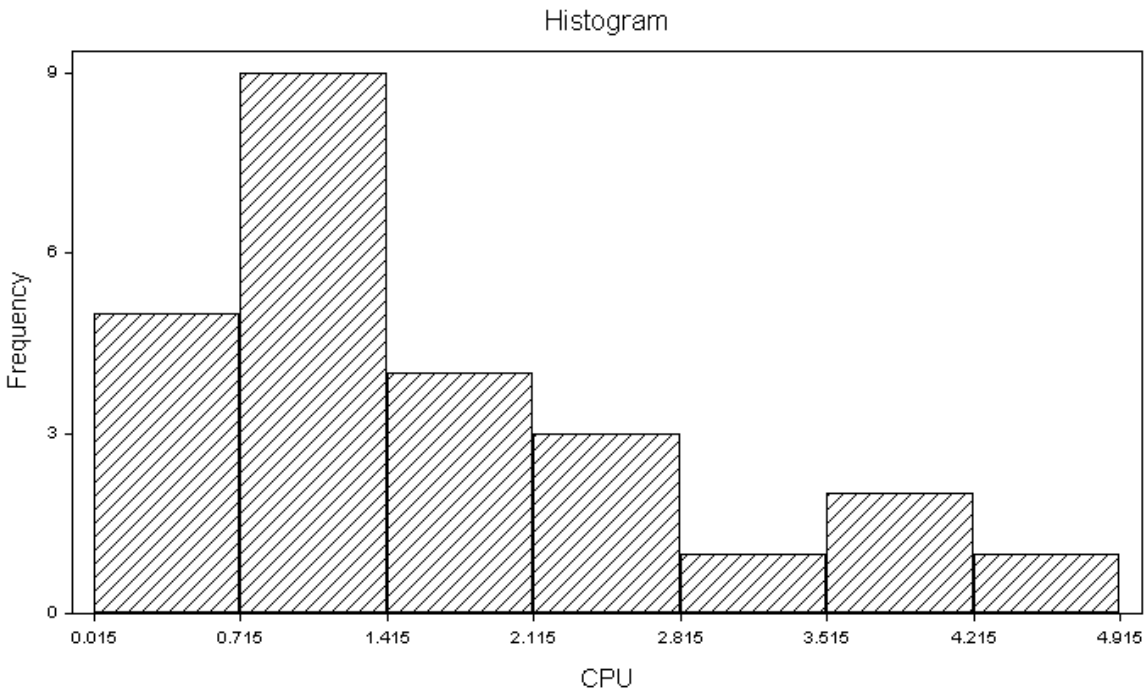
STEM AND LEAF PLOT OF CPU

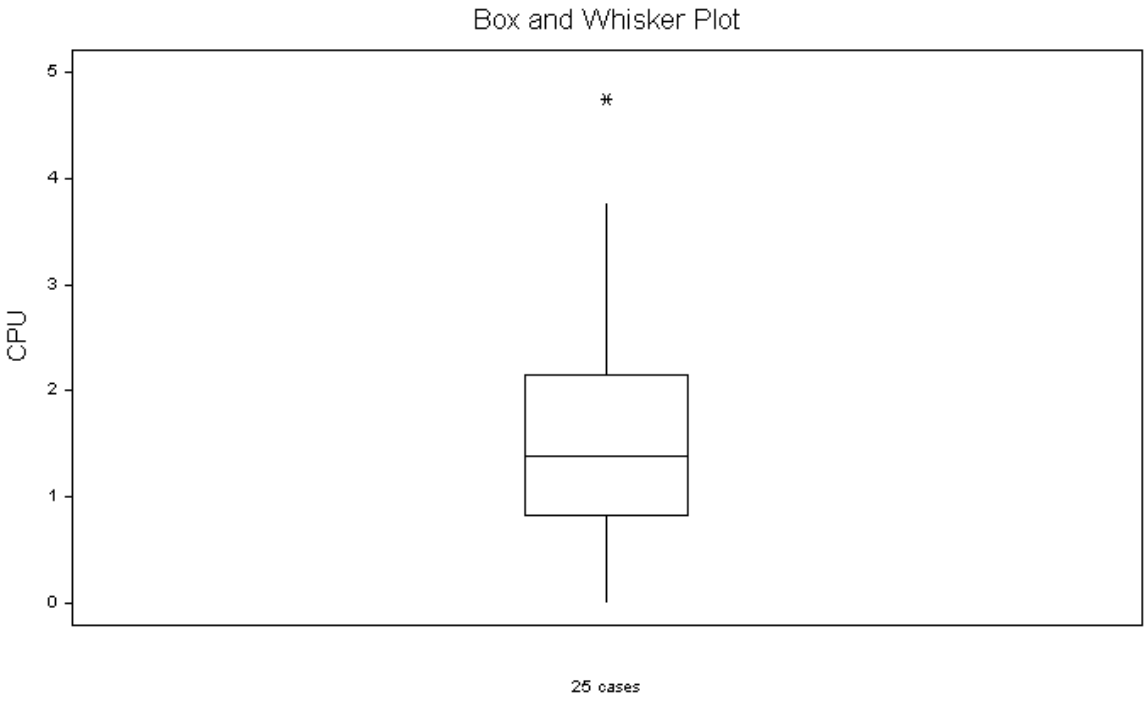
```
   LEAF DIGIT UNIT = 0.1          MINIMUM  0.0200
   4  7  REPRESENTS 4.7           MEDIAN   1.3800
                                  MAXIMUM  4.7500
         STEM  LEAVES
       4    0  0114
       9    0  77899
      (5)   1  11234
      11    1  569
       8    2  014
       5    2  5
       4    3  0
       3    3  57
       1    4
       1    4  7

25 CASES INCLUDED   0 MISSING CASES
```

DESCRIPTIVE STATISTICS

	CPU
N	25
MEAN	1.6300
SD	1.1928
MINIMUM	0.0200
1ST QUARTI	0.7850
MEDIAN	1.3800
3RD QUARTI	2.2850
MAXIMUM	4.7500
MAD	0.6300


Histogram
Frequency
9
6
3
0
0.015
0.715
1.415
2.115
2.815
3.515
4.215
4.915
CPU


Box and Whisker Plot
CPU
5
4
3
2
1
0
*
25 cases

Comparación de los tiempos de CPU de trabajos tomados al azar enviados a dos servers: