



Estadística I

Introducción - Contenido del curso

Kevin Pérez - Ingeniero de Sistemas - Estadístico, (E) Maestría en Ciencia de Datos.
Universidad Cooperativa de Colombia.

Capítulo I - Conceptos generales

- ¿Qué es Estadística?
- Conceptos Generales. (Población, Muestra, Muestreo, Estadígrafo, Parámetros, variable, dato, clasificación de la estadística, Etapas de una investigación estadística.)
- ¿Qué es variable?.
- Clasificación de variables

Capítulo II - Análisis y distribución de frecuencias

- Distribución de frecuencias.
- Análisis e interpretación de tablas con variables discretas y continuas.
- Medidas de tendencia central
- Medidas de posición
- Medidas de dispersión
- Medidas de forma y asimetría

Capítulo III - Análisis y Representaciones gráficas.

- Histograma de frecuencias.
- Polígonos de frecuencia y Ojivas.
- Diagramas lineales y pictogramas
- Diagramas de barras.
- Diagramas circulares o de pastel.
- Otros gráficos

Capítulo I - Orígenes de la Estadística

- Historia de la Estadística

Capítulo I - Orígenes de la Estadística

- Los orígenes históricos de la Estadística (descriptiva) hay que buscarlos en los procesos de recogida de datos, censos y registros sistemáticos, asumiendo un papel asimilable a una aritmética estatal para asistir al gobernante, que necesitaba conocer la riqueza y el número de sus súbditos con fines tributarios y políticos.
- Los primeros registros de riqueza y población que se conocen se deben a los egipcios. Ramsés II en el 1400 a.C. realizó el primer censo conocido de las tierras de Egipto, no siendo éste, se supone, ni el primero ni el último que se hiciera en las tierras bañadas por el Nilo.
- Posteriormente, desde el siglo III a.C., en las civilizaciones china y romana se llevan a cabo censos e inventarios de posesiones, que pueden considerarse precedentes institucionalizados de la recogida de datos demográficos y económicos de los Estados Modernos.

Capítulo I - Orígenes de la Estadística

- Hay que realizar una mención especial del período helénico, en el que las escuelas matemáticas se suceden. Centros como el de Quios, donde estudió Hipócrates (Hipócrates de Quios) el matemático, considerado como el inventor del método matemático y escuelas como las de Cirene, Megara y al final Atenas, donde se reúnen los matemáticos, unos alrededor de Protágoras y otros en torno a Sócrates.
- En la Edad Media se vuelve a la utilización de la Aritmética para la recogida de datos, existiendo menos interés por la elucubración matemática abstracta. Es en este período de tiempo cuando Carlomagno ordenó en su *“Capitulare de villis”* la creación de un registro de todos sus dominios y bienes privados.
- En el siglo XVII se producen avances sustanciales, y así, en las universidades alemanas se imparten enseñanzas de “Aritmética Política”, término con el que se designa la descripción numérica de hechos de interés para la Administración Pública. Destacados autores de Aritmética Política fueron los ingleses Graunt (1620-1674) y Petty (1623-1687).

Capítulo I - Orígenes de la Estadística

Con métodos de estimación en los que cabía la conjetura, la experimentación y la deducción, Graunt llega a estimar tasas de mortalidad para la población londinense, analizando además la verosimilitud de la información de que disponía. Por su parte, Petty, cuyas aportaciones estadísticas fueron menos relevantes.

Capítulo I - Conceptos Generales

- ¿Qué es Estadística?

Capítulo I - Conceptos Generales

La **Estadística** se ocupa del manejo de la información que pueda ser cuantificada. Implica esto la descripción de conjuntos de datos y la inferencia a partir de la información recolectada de un fenómeno de interés. La función principal de la estadística abarca: Resumir, Simplificar, Comparar, Relacionar, Proyectar. Entre las tareas que debe enfrentar un estudio estadístico están:

- Delimitar con precisión la población de referencia o el conjunto de datos en estudio, las unidades que deben ser observadas, las características o variables que serán medidas u observadas.
- Estrategias de Observación: Censo, Muestreo, Diseño de Experimental
- Recolección y Registro de la información

Capítulo I - Conceptos Generales

- Depuración de la información.
- Construcción de Tablas.
- Análisis Estadístico:
 - Producción de resúmenes gráficos y numéricos.
 - Interpretación de resultados.

Cuando los datos comprenden toda la población de referencia, hablamos de un Censo y cuando solo comprometen una parte de ella, hablamos de una muestra. En ambos casos es pertinente un análisis Descriptivo. En el segundo caso un análisis Inferencial.

Capítulo I - Conceptos Generales

A grandes rasgos podemos decir que una **Población** es el conjunto de toda posible información, o de los objetos, que permite estudiar un fenómeno de interés.

- Una **muestra** es un subconjunto de información representativa de una población.
- Las **Variables** resultan ser aquellas características de interés que desean ser medidas sobre los objetos o individuos seleccionados. En la mayoría de los casos lo que se pretende es estimar, a partir de la información recolectada de una muestra, características desconocidas de los objetos en dicha población de interés.
- Las características desconocidas de una población serán llamadas **Parámetros**. Las características calculadas a partir de una muestra son llamadas **Estadísticas**. Una Inferencia es una generalización obtenida a partir de una muestra aleatoria.

Capítulo I - Conceptos Generales

La Estadística puede dividirse en dos grandes ramas: Estadística Descriptiva y Estadística Inferencial.

- ***Estadística descriptiva:*** Es el conjunto de métodos usados para la organización y presentación (descripción) de la información recolectada. La información recolectada puede ser catalogada de dos maneras: Datos Cualitativos y Cuantitativos.
- ***Estadística inferencial:*** Comprende los métodos y procedimientos para deducir propiedades (hacer inferencias) de una población, a partir de una pequeña parte de la misma (muestra).

Capítulo I - Variables y atributos

Una primera clasificación del tipo de datos en función de que las observaciones resultantes del experimento sean de tipo cualitativo o cuantitativo, en el primero de los casos se tiene un atributo y en el segundo una variable. Para hacer referencia genéricamente a una variable o a un atributo se utilizará el término carácter.

- **Ejemplo 1:** Como ejemplos de **atributos** pueden considerarse el color del pelo de un colectivo de personas, su raza o el idioma que hablan y como **variables** su estatura, peso o edad.
- Para poder operar con un atributo es necesario asignar a cada una de sus clases un valor numérico, con lo que se transforma en una variable, esta asignación se hará de forma que los resultados que se obtengan al final del estudio sean fácilmente interpretables.

Capítulo I - Variables discretas y continuas

Dentro del conjunto de las variables se distingue entre **discretas** y **continuas**. Se dice que una variable es discreta cuando entre dos valores consecutivos no toma valores intermedios y que es continua cuando puede tomar cualquier valor dentro de un intervalo.

- **Ejemplo 2:** La **estatura** de un grupo de personas sería una variable **continua**, mientras que el **número de cabellos** que tienen en la cabeza sería una variable **discreta**.
- En la práctica todas las variables son discretas debido a la limitación de los aparatos de medida, y así, en el ejemplo de las estaturas, quizás se podría detectar una diferencia de una cienmilésima de metro, o a lo más, de una millonésima, pero dados dos individuos que se diferencien en una millonésima no puede detectarse otro que tenga una estatura intermedia. De todas formas, en general se trata a las variables “*teóricamente*” continuas como tales, por razones que se pondrán de manifiesto más adelante.

Capítulo I - Tipos de variables y escalas de medida

Se denomina escalamiento al desarrollo de reglas sistemáticas y de unidades significativas de medida para identificar o cuantificar las observaciones empíricas. La clasificación más común distingue cuatro conjuntos de reglas básicas que producen cuatro escalas de medida; éstas son:

- La escala de medida más simple implica una relación de identidad entre el sistema de números y el sistema empírico objeto de medida. La escala resultante se denomina **nominal**, porque los números empleados se consideran como “etiquetas” las cuales se asignan a los objetos con el propósito de clasificarlos, pero no poseen el significado numérico usual, aparte de la relación de igualdad; por tanto, tienen una naturaleza no métrica. *El género, la raza, la profesión, el credo religioso, son variables observadas en este tipo de escala.*

Capítulo I - Tipos de variables y escalas de medida

- Una escala más compleja, implica además de la relación de igualdad como el caso anterior, una relación de orden que se preserva tanto en el sistema numérico como en el sistema empírico (medidas sobre los objetos). Éste tipo de escalas se denomina **ordinal** porque los números que se asignan a los atributos deben respetar (conservar) el orden de la característica que se mide. El tipo de datos que resulta tiene naturaleza no métrica. La valoración de la opinión en “*de acuerdo*”, “*indiferente*” o “*en desacuerdo*”, constituye un ejemplo de una variable típica de esta escala.

Capítulo I - Tipos de variables y escalas de medida

- El siguiente nivel de escalamiento implica, además de una relación de orden como la escala anterior, una relación de igualdad de diferencias entre pares de objetos respecto a una característica determinada. La escala resultante se denomina de **intervalo** porque las diferencias entre los números se corresponden con las diferencias entre la propiedad medida sobre los objetos, y por tanto tiene naturaleza métrica. La medición de la *temperatura, la altura física, constituyen ejemplos de esta escala de medida*. Una característica adicional de esta escala es la necesidad de precisar un origen o punto “cero” respecto al cual la medida tiene sentido, esto no necesariamente significa ausencia del atributo. En el ejemplo de la temperatura, el cero en la escala Celsius, es la temperatura de congelación del agua al nivel del mar; nótese que este cero no corresponde con el de la escala Fahrenheit.

Capítulo I - Tipos de variables y escalas de medida

- El nivel más complejo de escalamiento implica, además de una relación de igualdad de diferencias como en la escala anterior, un punto de origen fijo o natural, el cero absoluto. El resultado es la escala de **razón**, que tiene también naturaleza métrica. Ejemplos de este tipo de escala son el peso, la talla o la edad de los individuos.

Capítulo I - Resumen

- Las variables pueden ser **cuantitativas**, cuando se realiza una medición y el resultado es un número, o pueden ser **cualitativas**, cuando solamente registran una cualidad o atributo del objeto o persona en estudio. *La edad, el peso y la estatura son ejemplos de variables cuantitativas en una población de personas, mientras que el sexo y el estado civil son variables cualitativas.*
- De acuerdo al número total de sus posibles valores, una variable cuantitativa puede ser clasificada como discreta cuando sólo puede tomar un número discreto (es decir, finito o numerable) de valores, o continua cuando puede tomar cualquier valor dentro de un intervalo (a, b) de la recta real.

Capítulo I - Resumen

De acuerdo con la posible relación que pudieran guardar los valores de una variable, se cuenta por lo menos con cuatro escalas de medición. Las variables cualitativas pueden ser clasificadas de acuerdo a dos escalas: escala nominal o escala ordinal. Mientras que las variables **cuantitativas** pueden clasificarse por: **escala de intervalo** o **escala de razón**.

Capítulo I - Distribuciones de frecuencia

La organización de los datos constituye la primera etapa de su tratamiento, pues, facilita los cálculos posteriores y evita posibles confusiones. La organización va a depender del número de observaciones distintas que se tengan y de las veces que se repitan cada una de ellas. En base a lo anterior se pueden estructurar los datos de tres maneras distintas:

- **Tipo I:** Cuando se tiene un número pequeño de observaciones casi todas distintas, éstas se darán por extensión.
- **Ejemplo 1:** En la serie: 2, 3, 5, 7, 7, 8, 11, 14, 16, 19, el 7 se repite dos veces y el resto de los valores está presente una vez.
- **Ejemplo 2:** En la serie: -5, 2, -3, -2, 4, 4, 2, 5, el 2 y el 4 se repiten dos veces y el resto de los valores está presente una vez.
- **Ejercicio** ¿Que pasa en la siguiente serie? 3.5, 2, 7.5, 3.5, 1.5, 4.5, 7.5.

Capitulo I - Distribuciones de frecuencia

Tipo II: Cuando se tiene un gran número de observaciones pero muy pocas distintas, se organizan en una tabla de frecuencias, es decir, cada uno de los valores acompañado de la frecuencia con la que se presenta.

La tabla

<i>Valor</i>	<i>Frecuencia</i>
2	4
4	4
5	3
6	2
7	3
8	3
9	1

Indica que el valor 2 se repite 4 veces, el valor 4 se repite 4 veces, el valor 5 se repite 3 veces, etc ...

Capítulo I - Distribuciones de frecuencias

Para efectuar cálculos, sea cuál sea el tipo de distribución, se disponen los datos de la siguiente forma:

x_i	f_i	F_i	h_i	H_i
x_1	f_1	F_1	h_1	H_1
x_2	f_2	F_2	h_2	H_2
\vdots	\vdots	\vdots	\vdots	\vdots
x_r	f_r	$N_r = n$	f_r	$H_r = 1$

Donde:

- n representa al número total de observaciones (datos) en la muestra
- x_i : variable es la variable de interés.
- f_i es la frecuencia absoluta, definida como la cantidad de veces que se repite la observación (dato) en la muestra.

Capítulo I - Distribuciones de frecuencias

- F_i es la frecuencia absoluta acumulada, que se obtiene como la suma acumulada de las frecuencias absolutas, $\sum_{i=1}^r f_i$
- h_i es la frecuencia relativa, definida como el cociente (división) de las frecuencias absolutas entre el tamaño de la muestra, $\frac{f_i}{n}$
- H_i es la frecuencia relativa acumulada, que viene dada por la suma acumulada de las frecuencias relativas acumuladas, $\sum_{i=1}^r f_i$
- **Ejemplo 1:** Durante el mes de Enero, en una ciudad de Montería, se han registrado las siguientes temperaturas máximas: 32, 31, 28, 29, 33, 32, 31, 30, 31, 31, 27, 28, 29, 30, 32, 31, 31, 30, 30, 29, 29, 30, 30, 31, 30, 31, 34, 33, 33, 29, 29. Construir la tabla de frecuencias.

Capítulo I - Distribuciones de frecuencias

Se procede entonces según el siguiente procedimiento.

1. Se identifican los elementos que se repiten en la muestra dada.
2. Se realizan los conteos de las repeticiones de cada elemento de la muestra (Frecuencia absoluta).
3. Se calcula la Frecuencia absoluta acumulada según la formula
4. Se calcula la Frecuencia relativa según la formula
5. Se calcula la Frecuencia relativa acumulada según la formula

Capítulo I - Distribuciones de frecuencias

- Se identifican los elementos que se repiten en la muestra dada.

x_i	f_i	F_i	h_i	H_i
27				
28				
29				
30				
31				
32				
33				
34				

Capítulo I - Distribuciones de frecuencias

- Se calculan las Frecuencias absolutas.

x_i	f_i	F_i	h_i	H_i
27	1			
28	2			
29	6			
30	7			
31	8			
32	3			
33	3			
34	1			

Capítulo I - Distribuciones de frecuencias

- Se calculan las Frecuencias absolutas acumuladas.

x_i	f_i	F_i	h_i	H_i
27	1	1		
28	2	3		
29	6	9		
30	7	16		
31	8	24		
32	3	27		
33	3	30		
34	1	31		

Capítulo I - Distribuciones de frecuencias

- Se calculan las Frecuencias relativas.

x_i	f_i	F_i	h_i	H_i
27	1	1	0.032	
28	2	3	0.065	
29	6	9	0.194	
30	7	16	0.226	
31	8	24	0.258	
32	3	27	0.097	
33	3	30	0.097	
34	1	31	0.032	

Capítulo I - Distribuciones de frecuencias

- Se calculan las Frecuencias relativas acumuladas.

x_i	f_i	F_i	h_i	H_i
27	1	1	0.032	0.032
28	2	3	0.065	0.097
29	6	9	0.194	0.290
30	7	16	0.226	0.516
31	8	24	0.258	0.774
32	3	27	0.097	0.871
33	3	30	0.097	0.968
34	1	31	0.032	1

Capítulo I - Distribuciones de frecuencias

Ejemplo 2: El número de veces que han ido al cine en el último mes los alumnos de una clase es: 2, 3, 0, 1, 5, 3, 2, 1, 2, 3, 5, 0, 5, 4, 1, 1, 1, 2, 0, 1, 2. Construir la tabla de frecuencias.

Capítulo I - Distribuciones de frecuencias

- Se identifican los elementos que se repiten en la muestra dada.

x_i	f_i	F_i	h_i	H_i
0				
1				
2				
3				
4				
5				

Capítulo I - Distribuciones de frecuencias

- Se calculan las frecuencias absolutas.

x_i	f_i	F_i	h_i	H_i
0	5			
1	7			
2	6			
3	3			
4	1			
5	3			

Capítulo I - Distribuciones de frecuencias

- Se calculan las frecuencias absolutas acumuladas.

x_i	f_i	F_i	h_i	H_i
0	5	5		
1	7	12		
2	6	18		
3	3	21		
4	1	22		
5	3	25		

Capítulo I - Distribuciones de frecuencias

- Se calculan las frecuencias relativas.

x_i	f_i	F_i	h_i	H_i
0	5	5	0.20	
1	7	12	0.28	
2	6	18	0.24	
3	3	21	0.12	
4	1	22	0.04	
5	3	25	0.12	

Capítulo I - Distribuciones de frecuencias

- Se calculan las frecuencias relativas.

x_i	f_i	F_i	h_i	H_i
0	5	5	0.20	0.20
1	7	12	0.28	0.48
2	6	18	0.24	0.72
3	3	21	0.12	0.84
4	1	22	0.04	0.88
5	3	25	0.12	1

Capítulo I - Distribuciones de frecuencias

- **Ejercicio 1.** Las notas de un examen de matemáticas de 30 alumnos de una clase son las siguientes: 5, 3, 4, 1, 2, 8, 9, 8, 7, 6, 6, 7, 9, 8, 7, 7, 1, 0, 1, 5, 9, 9, 8, 0, 8, 8, 8, 9, 5, 7. Se pide construir la tabla de frecuencias
- **Ejercicio 2.** El número de faltas de ortografía que cometieron un grupo de estudiantes en un dictado fue: 0, 3, 1, 2, 0, 2, 1, 3, 0, 4, 0, 1, 1, 4, 3, 5, 3, 2, 4, 1, 5, 0, 2, 1, 0, 0, 0, 0, 2, 1, 2, 1, 0, 0, 3, 0, 5, 3, 2, 1. Se pide construir la tabla de frecuencias

Capítulo I - Distribuciones de frecuencias

Tipo III: En el caso de que haya muchas observaciones, la mayoría de ellas distintas, pueden disponerse agrupándolas en intervalos e indicando el número de observaciones que caen dentro de cada intervalo.

Capítulo I - Distribuciones de frecuencias

Ejemplo 1: La tabla

<i>Intervalo</i>	<i>Frecuencia</i>
[2, 3)	4
[3, 7)	6
[7, 12)	12
[12, 21)	8
[21, 25)	6
[25, 30)	4
[30, 50)	3

Nos dice que en el intervalo (2, 3] hay 4 observaciones, que en el (3, 7] hay 6, etc ...

Capítulo I - Distribuciones de frecuencias

Para efectuar cálculos, en este tipo de distribución, se disponen los datos de la siguiente forma:

<i>Clases</i>	f_i	F_i	h_i	H_i	x_i
$L_{I_1} - L_{S_1}$	f_1	F_1	h_1	H_1	x_1
$L_{I_2} - L_{S_2}$	f_2	F_2	h_2	H_2	x_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$L_{I_r} - L_{S_r}$	f_r	$N_r = n$	f_r	$H_r = 1$	x_r

Donde:

- n representa al número total de observaciones (datos) en la muestra
- L_{I_i} Es el límite inferior de la clase
- L_{S_i} Es el límite superior de la clase

Capítulo I - Distribuciones de frecuencias

- f_i es la frecuencia absoluta, definida como la cantidad de veces que se repite la observación (dato) en la muestra.
- F_i es la frecuencia absoluta acumulada, que se obtiene como la suma acumulada de las frecuencias absolutas, $\sum_{i=1}^r f_i$
- h_i es la frecuencia relativa, definida como el cociente (división) de las frecuencias absolutas entre el tamaño de la muestra, $\frac{f_i}{n}$
- H_i es la frecuencia relativa acumulada, que viene dada por la suma acumulada de las frecuencias relativas acumuladas, $\sum_{i=1}^r f_i$
- x_i : es la marca de clases y se calcula como el promedio de los límites de las clases.

Capítulo I - Distribuciones de frecuencias

El procedimiento para la consecución de el número de clases con sus respectivos límites es el siguiente:

1. **Calcule el Rango (R)** También se llama recorrido o amplitud total. Es la diferencia entre el valor mayor y el menor de los datos.
2. Seleccione el número de intervalos de **Clase k** . Para calcular el número de intervalos se aplica la regla de Sturges:

$$k = 1 + 3.32 \log(n)$$

3. **Calcule el Ancho del Intervalo (C_i)**. Se obtiene dividiendo el Rango por el número de intervalos

$$C_i = \frac{R}{K}$$

4. Forme los intervalos de clase agregando $C_i - UMP$ al límite inferior de cada clase, comenzando por el mínimo de la distribución.

Capítulo I - Distribuciones de frecuencias

Ejemplo 3: Los siguientes datos corresponden al número de clientes que acudieron a un cafe internet durante 40 días.

30, 35, 15, 21, 18, 32, 39, 20, 19, 20, 20, 34, 13, 13, 20, 35, 30, 17, 30, 31, 10, 32, 22, 14, 30, 36, 23, 14, 20, 34, 20, 28, 20, 16, 31, 38, 22, 12, 28, 13

Para elaborar la tabla de distribución de frecuencia, seguir los siguientes pasos.

- Paso 1. Ordenar de menor a mayor.

10 12 13 13 13 14 14 15 16 17 18 19 20 20 20 20 20 20 20 21 22 22 23 28 28 30 30 30 30 31 31 32 32 34 34 35 35 36 38 39.

- Paso 2. Calcular el rango

$$R = X_{\text{máx}} - X_{\text{mín}} = 39 - 10 = 29$$

Capítulo I - Distribuciones de frecuencias

- Paso 3. Calcular el número de clases y la $UMP/2$

$$k = 1 + 3.32 \log(40) = 6.3 \cong 6, \quad UMP = 1, \quad \text{Luego} \quad UMP/2 = 0.5$$

- Paso 4. Calcular el ancho de clase.

$$C_i = \frac{R}{k} = \frac{29}{6} = 4.8 \cong 5$$

- Paso 5. Formar los intervalos de clase agregando $C_i - UMP$ al límite inferior de cada clase

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>
1	$Li_1 = X_{\text{mín}}$	$Li_1 + C_i - UMP$	$Li_1 - UMP/2$	$Li_1 + C_i - UMP/2$
2	$Li_2 = Li_1 + C_i$	$Li_2 + C_i - UMP$	$Li_2 - UMP/2$	$Li_2 + C_i - UMP/2$
3	$Li_3 = Li_2 + C_i$	$Li_3 + C_i - UMP$	$Li_3 - UMP/2$	$Li_3 + C_i - UMP/2$
4	$Li_4 = Li_3 + C_i$	$Li_4 + C_i - UMP$	$Li_4 - UMP/2$	$Li_4 + C_i - UMP/2$
5	$Li_5 = Li_4 + C_i$	$Li_5 + C_i - UMP$	$Li_5 - UMP/2$	$Li_5 + C_i - UMP/2$
6	$Li_6 = Li_5 + C_i$	$Li_6 + C_i - UMP$	$Li_6 - UMP/2$	$Li_6 + C_i - UMP/2$

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>
1	10	$Ls_1 = Li_1 + C_i - UMP$	$Lri_1 = Li_1 - UMP/2$	$Lrs_1 = Ls_1 + UMP/2$
2	15	$Ls_2 = Li_2 + C_i - UMP$	$Lri_2 = Li_2 - UMP/2$	$Lrs_2 = Ls_3 + UMP/2$
3	20	$Ls_3 = Li_3 + C_i - UMP$	$Lri_3 = Li_3 - UMP/2$	$Lrs_3 = Ls_3 + UMP/2$
4	25	$Ls_4 = Li_4 + C_i - UMP$	$Lri_4 = Li_4 - UMP/2$	$Lrs_4 = Ls_4 + UMP/2$
5	30	$Ls_5 = Li_5 + C_i - UMP$	$Lri_5 = Li_5 - UMP/2$	$Lrs_5 = Ls_5 + UMP/2$
6	35	$Ls_6 = Li_6 + C_i - UMP$	$Lri_6 = Li_6 - UMP/2$	$Lrs_6 = Ls_6 + UMP/2$

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>
1	10	14	$Lri_1 = Li_1 - UMP/2$	$Lrs_1 = Ls_1 + UMP/2$
2	15	19	$Lri_2 = Li_2 - UMP/2$	$Lrs_2 = Ls_3 + UMP/2$
3	20	24	$Lri_3 = Li_3 - UMP/2$	$Lrs_3 = Ls_3 + UMP/2$
4	25	29	$Lri_4 = Li_4 - UMP/2$	$Lrs_4 = Ls_4 + UMP/2$
5	30	34	$Lri_5 = Li_5 - UMP/2$	$Lrs_5 = Ls_5 + UMP/2$
6	35	39	$Lri_6 = Li_6 - UMP/2$	$Lrs_6 = Ls_6 + UMP/2$

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>
1	10	14	9.5	$Lrs_1 = Ls_1 + UMP/2$
2	15	19	14.5	$Lrs_2 = Ls_3 + UMP/2$
3	20	24	19.5	$Lrs_3 = Ls_3 + UMP/2$
4	25	29	24.5	$Lrs_4 = Ls_4 + UMP/2$
5	30	34	29.5	$Lrs_5 = Ls_5 + UMP/2$
6	35	39	34.5	$Lrs_6 = Ls_6 + UMP/2$

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>
1	10	14	9.5	14.5
2	15	19	14.5	19.5
3	20	24	19.5	24.5
4	25	29	24.5	29.5
5	30	34	29.5	34.5
6	35	39	34.5	39.5

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>	<i>f_i</i>
1	10	14	9.5	14.5	7
2	15	19	14.5	19.5	5
3	20	24	19.5	24.5	11
4	25	29	24.5	29.5	2
5	30	34	29.5	34.5	10
6	35	39	34.5	39.5	5

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>	<i>f_i</i>	<i>F_i</i>
1	10	14	9.5	14.5	7	7
2	15	19	14.5	19.5	5	12
3	20	24	19.5	24.5	11	23
4	25	29	24.5	29.5	2	25
5	30	34	29.5	34.5	10	35
6	35	39	34.5	39.5	5	40

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>	f_i	F_i	h_i
1	10	14	9.5	14.5	7	7	0.175
2	15	19	14.5	19.5	5	12	0.125
3	20	24	19.5	24.5	11	23	0.275
4	25	29	24.5	29.5	2	25	0.05
5	30	34	29.5	34.5	10	35	0.25
6	35	39	34.5	39.5	5	40	0.125

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>	f_i	F_i	h_i	H_i
1	10	14	9.5	14.5	7	7	0.175	0.175
2	15	19	14.5	19.5	5	12	0.125	0.3
3	20	24	19.5	24.5	11	23	0.275	0.575
4	25	29	24.5	29.5	2	25	0.05	0.625
5	30	34	29.5	34.5	10	35	0.25	0.875
6	35	39	34.5	39.5	5	40	0.125	1

Capítulo I - Distribuciones de frecuencias

<i>Clase</i>	<i>Li</i>	<i>Ls</i>	<i>Lri</i>	<i>Lrs</i>	f_i	F_i	h_i	H_i	x_i
1	10	14	9.5	14.5	7	7	0.175	0.175	12
2	15	19	14.5	19.5	5	12	0.125	0.3	17
3	20	24	19.5	24.5	11	23	0.275	0.575	22
4	25	29	24.5	29.5	2	25	0.05	0.625	27
5	30	34	29.5	34.5	10	35	0.25	0.875	32
6	35	39	34.5	39.5	5	40	0.125	1	37

Capítulo I - Distribuciones de frecuencia

Ejercicio 1: Los siguientes datos corresponden al número de goles anotados en 40 partidos de futbol sala en una liga local.

3, 8, 2, 3, 8, 4, 10, 15, 4, 5, 7, 12, 6, 1, 9, 6, 4, 4, 7, 13, 8, 18, 9, 2, 11, 9, 7, 2, 14, 2, 1, 24, 10, 4, 2, 7, 6, 6, 1, 8. Se pide calcular la tabla de frecuencia por intervalos.

Ejercicio 2: Los siguientes datos se refieren al diámetro en pulgadas de un engrane. 6.75, 7.00, 7.00, 6.75, 6.50, 6.50, 7.15, 7.00, 6.50, 6.50, 6.50, 6.25, 6.25, 6.50, 6.65, 7.00, 7.25, 6.70, 6.00, 6.75, 6.00, 6.75, 6.75, 7.10, 7.00, 6.70, 6.50, 6.75, 6.25, 6.65, 6.75, 7.10, 7.25, 6.75, 6.25, 6.25, 7.00, 6.75, 7.00, 7.15 Se pide calcular la tabla de frecuencia por intervalos.

Capítulo I - Distribuciones de frecuencia - Taller No. 1

Ejercicio 1: Se ha realizado una encuesta en 30 hogares en la que se les pregunta el nº de individuos que conviven en el domicilio habitualmente. Las respuestas obtenidas han sido las siguientes: 4, 4, 1, 3, 5, 3, 2, 4, 1, 6, 2, 3, 4, 5, 5, 6, 2, 3, 3, 2, 2, 1, 8, 3, 5, 3, 4, 7, 2, 3.

- Calcule la distribución de frecuencias de la variable obteniendo las frecuencias absolutas, relativas y sus correspondientes acumuladas.
- ¿Qué porcentaje de hogares está compuesto por tres o menos personas?

Ejercicio 2: El número de personas que se presentaron al examen para el permiso de conducir en 40 convocatorias fue: 58, 66, 60, 72, 66, 70, 81, 59, 73, 68, 52, 75, 73, 68, 75, 78, 65, 68, 63, 59, 74, 80, 79, 75, 76, 66, 68, 76, 72, 70, 67, 63, 68, 80, 74, 64, 65, 79, 54, 59. Se pide realizar la tabla de frecuencia por intervalos.

Capítulo I - Medidas de Tendencia Central

Una vez organizados los datos en su correspondiente distribución de frecuencias, se procede a dar una serie de medidas que resuman toda esa información y que, “de alguna manera”, representen a la distribución.

- **La Media Aritmética:** La media es una medida de representación central que se define para cada uno de los tipos de distribución como sigue:
 - Distribución Tipo I:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Capítulo I - Medidas de Tendencia Central

- Distribución Tipo II:

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{n}$$

Donde x_i son los valores que toma la variable y f_i son las frecuencias absolutas.

- Distribución Tipo III:

$$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{n}$$

Donde x_i son las denominadas marcas de clase y f_i son las frecuencias absolutas.

Capítulo I - Medidas de Tendencia Central

Ejemplo Tipo I: La media de la distribución del ejemplo 1 viene dada por:

$$\bar{x} = \frac{(2 + 3 + 5 + 7 + 7 + 8 + 11 + 14 + 16 + 19 + 7)}{11} = \frac{99}{11} = 9$$

Ejemplo Tipo II: Para los datos de la tabla ejemplo en la distribución tipo II tenemos que:

$$\bar{x} = \frac{(2 \cdot 4 + 4 \cdot 4 + 5 \cdot 3 + \dots + 9 \cdot 1)}{20} = \frac{105}{20} = 25.5$$

Ejemplo Tipo III: Para los datos de la tabla del ejemplo 3 en la distribución tipo III tenemos que:

$$\bar{x} = \frac{(7 \cdot 12 + 5 \cdot 17 + 22 \cdot 11 + \dots + 5 \cdot 37)}{40} = \frac{970}{40} = 24.5$$

