



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
FACULTAD DE CONTADURÍA Y ADMINISTRACIÓN

## TABLA DE DISTRIBUCIÓN DE FRECUENCIAS

L.C. y Mtro. Francisco Javier Cruz Ariza

Una manera de simplificar los datos es usar una tabla de frecuencia o una distribución de frecuencia. La distribución de frecuencias es una tabla que organiza los datos en clases, es decir en grupos de valores que describen una característica de los datos.

Te recomiendo que previamente, copies los datos a tabular en una hoja de Excel. Es importante que los datos queden en una sola columna, y que estén ordenados de menor a mayor, para facilitar el trabajo.

## CONSTRUCCIÓN DE LA TABLA

**PASO No. 1. Determinar el rango:** (Diferencia entre el dato mayor y el menor)

$$\text{Rango} = \text{Valor máximo} - \text{valor mínimo}$$

Para que haya congruencia, se recomienda que el rango contenga el mismo número de decimales que los datos de origen. *Ejemplo:* si el dato mayor es 9.1 y el menor 2.5, el rango es 6.6

**PASO No. 2. Determinar el número de intervalos de clase:** Significa en cuántas categorías o subgrupos vamos a clasificar o agrupar nuestros datos. Para determinar el número óptimo de intervalos de clase, en los cuales nuestros datos quedarán perfectamente distribuidos, aplicamos la Regla de Sturges:

$$\text{Regla de Sturges} \rightarrow \text{No. de intervalos de clase} = 1 + (3.322 * (\log n))$$

En donde “n” representa el número total de datos u observaciones que tenemos recopilados.

Evidentemente, el número de intervalos debe ser **exacto**; es decir, un número entero.

**PASO No. 3. Cálculo de la Amplitud:** Ahora procederemos a calcular el ancho de los intervalos de clase; es decir, la dimensión que tendrá cada uno de los intervalos que determinamos en el paso anterior.

$$\text{Amplitud} = \frac{\text{Rango}}{\text{No. de intervalos de Clase}}$$

Es importante comentar que para el cálculo de la amplitud se tiene que considerar que la amplitud tiene que tener, exactamente, el mismo número de decimales que manejan los datos originales. Supongamos, por ejemplo, que nuestros datos tienen 2 decimales; la amplitud tendrá que ser expresada también en dos decimales; por decir, 2.85

Para validar si los parámetros que establecimos están bien calculados, debemos proceder a aplicar los siguientes cálculos. (Se recomienda introducir una función condicional en Excel):

## REGLAS QUE DEBEN CUMPLIRSE:

**Regla No. 1:** No. de Intervalos de clase x Amplitud > Rango

**Regla No. 2:** No. de Intervalos de clase x Amplitud - Rango  $\leq$  Amplitud

Es **indispensable** que se cumplan las dos reglas a la perfección, ya que de no hacerlo, la tabla será construida con cierto grado de error. Para ello, podemos ajustar el número de intervalos de clase o la amplitud a la alza o a la baja, según se necesite.

### Ejemplo:

Pensemos que el cálculo para determinar el número de intervalos de clase fue 8.3. Como no podemos construir fracciones de intervalos, sino enteros, primero probamos redondeando a la baja este número a 8. Si la regla se cumple satisfactoriamente, proseguimos. En caso de que no se cumpla, lo que correspondería, en primer instancia, es subir el número de intervalos a 9 (aunque se vea muy disparatado) y probar lo que sucede.

Si tampoco este ajuste resultó suficiente, debemos proceder a ajustar la amplitud. Considerando que vamos a redondear el número de intervalos a 8 (según el caso anterior) y que la amplitud fuera 3.783, sabemos que ésta debe considerarse con el mismo número de decimales que los datos de estudio, así que probaremos primero ajustarla en 3.78 y ver qué sucede. Si se cumplen las reglas considerando 8 intervalos y una amplitud de 3.78, proseguimos; si no, podemos considerar 8 intervalos y amplitud de 3.79, o experimentar con 9 intervalos y amplitud de 3.78, etcétera. El objetivo es que se **deben** cumplir las reglas.

### Encabezado de la Tabla de Distribución de Frecuencias:

(Copiar directamente formato en Excel)

Nº Intervalo	Límites de Clase		Límites Reales de Clase		Marca de Clase	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Acumulada	Frecuencia Acumulada %	Frecuencia Complementaria	Frecuencia Complementaria %
	L.I.C.	L.S.C.	L.R.I.C.	L.R.S.C.	$X_i$	$f_i$	$f_i \%$	"- que"	"- que %"	" + que"	" + que"

### PASO No. 4: Tabulación de los datos:

#### 4.1. Límites de Clase (L.I.C. y L.S.C.):

Los límites de clase definen el ancho de cada intervalo; es decir, en qué número inicia nuestro intervalo (Límite Inferior de Clase, o L.I.C.), y hasta dónde llega (Límite Superior de Clase, o L.S.C.)

El límite inferior de la primera clase (es decir, el inicio del primer intervalo), siempre será el dato menor. La distancia entre el primer intervalo y el segundo, será siempre igual a la amplitud, y así sucesivamente.

*Supongamos la siguiente información:*

Dato menor: 2.86

Amplitud: 1.21

*Entonces:*

Límite Inferior de Clase del Intervalo 1: 2.86

Límite Inferior de Clase del Intervalo 2: 4.07 (2.86 + 1.21)

Límite Inferior de Clase del Intervalo 3: 5.28 (4.07 + 1.21)

La distancia entre todos y cada uno de los L.I.C. es la *Amplitud* (constante), que en este ejemplo es de 1.21.

**Se sugiere determinar el L.I.C. de todos los intervalos de la tabla, antes de proceder.**

Para establecer los Límites Superiores de Clase, se recomienda seguir la siguiente tabla:

Nº de decimales de los datos originales	L. S.C.
0 decimales (Enteros)	Restar "1" al L.I.C. inmediato posterior
1 decimal	Restar ".1" al L.I.C. inmediato posterior
2 decimales	Restar ".01" al L.I.C. inmediato posterior
3 decimales	Restar ".001" al L.I.C. inmediato posterior

*En consecuencia, siguiendo nuestro ejemplo:*

Límite Superior de Clase del Intervalo 1: 4.06 (4.07 – 0.01)

Límite Superior de Clase del Intervalo 2: 5.27 (5.28 – 0.01)

#### **4.2. Límites Reales de Clase (L.R.I.C. y L.R.S.C.):**

Son creados para evitar espacios o huecos en los datos, o bien para establecer un criterio de discriminación para que no exista duda en cuanto al lugar que debe ocupar cada uno de los datos, al categorizarlos o tabularlos en cada uno de los intervalos.

Supongamos, como ejemplo, y siguiendo los supuestos del caso anterior, que tenemos un dato con 3 decimales, siendo que los demás datos tienen solo 2. Por ejemplo, tenemos un dato cuyo valor numérico es 4.066. Como podemos apreciar, este dato pudiera ser ubicado tanto en el intervalo 1, cuyo *L.S.C.* es 4.06, como en el 2, cuyo *L.I.C.* es 4.07. Para saber dónde ubicarlo, nos son útiles los *Límites Reales de Clase*.

También existen límites reales inferiores de clase y límites reales superiores de clase. En su construcción se debe considerar la siguiente tabla:

Nº de decimales de los datos originales	L.R.I.C.	L.R.S.C.
0 decimales (Enteros)	- 0.5 (restar al L.I.C. del mismo intervalo)	0.5 (sumar al L.S.C. del mismo intervalo)
1 decimal	- 0.05	0.05
2 decimales	-0.005	0.005
3 decimales	-0.0005	0.0005

Retomando el ejemplo de los Límites de clase, tenemos:

Límite Inferior de Clase del Intervalo 1: 2.86

Límite Superior de Clase del Intervalo 1: 4.06

✚ Los datos originales tienen 2 decimales:

Nº de decimales de los datos originales	L.R.I.C.	L.R.S.C.
0 decimales (Enteros)	- 0.5 (restar al L.I.C. del mismo intervalo)	0.5 (sumar al L.S.C. del mismo intervalo)
1 decimal	- 0.05	0.05
<b>2 decimales</b>	<b>-0.005</b>	<b>0.005</b>
3 decimales	-0.0005	0.0005

$$\text{L.R.I.C. del intervalo 1} = \text{L.I.C.} - .005 = 2.86 - .005 = 2.855$$

$$\text{L.R.S.C. del intervalo 1} = \text{L.S.C.} + .005 = 4.06 + .005 = 4.065$$

✚ Procedemos con el mismo procedimiento para calcular el resto de los límites reales para los intervalos posteriores:

Límite Inferior de Clase del Intervalo 2: 4.07

Límite Superior de Clase del Intervalo 2: 5.27

$$\text{L.R.I.C. del intervalo 2} = \text{L.I.C.} - .005 = 4.07 - .005 = 4.065 \rightarrow \text{COINCIDE CON EL L.R.S.C. DEL}$$

INTERVALO ANTERIOR (NO DEJAMOS NINGÚN HUECO O ESPACIO EN BLANCO).

L.R.S.C. del intervalo 2 = L.S.C. + .005 = 5.27 + .005 = 5.275

Retomando la inquietud original que teníamos, de dónde ubicar al dato cuyo valor era de 4.066, por criterio establecemos que se encuentra más cercano al intervalo 2 que al 1.

#### 4.3. Marcas de Clase ( $X_i$ ):

La marca de clase de un intervalo es el punto medio del intervalo. Se obtiene sumando el límite inferior de clase mas el límite superior y esto dividiéndolo entre 2; también se puede obtener por los límites reales de clase y este será el límite real inferior más el límite real superior y esto dividiéndolo entre 2.

En Excel lo podemos calcular fácilmente mediante el promedio aritmético del L.I.C. y el L.S.C. Debe contener el mismo número de decimales que los datos originales.

La marca de clase toma el valor de la variable “Xi” y a partir de que se establece se convierte en el valor único del intervalo para toda la etapa del análisis.

#### 4.4. Frecuencia Absoluta ( $F_i$ ):

La frecuencia es el número de veces que se repite un evento, o bien, el número de datos que cae o queda contenido en cada intervalo de clase. La suma de la columna de la frecuencia es siempre igual al No. de datos del experimento o estudio.

Para obtenerla manualmente, se debe contar el número de datos que caen dentro de los límites de clase de cada intervalo.

Siguiendo el ejemplo anterior, sabemos que el L.I.C del intervalo 1 es 2.86, y el L.S.C. es 4.06. Lo que tenemos que hacer, es contar cuántos de nuestros datos se encuentran dentro de estos valores, y anotarlos en la columna correspondiente, y así sucesivamente hasta terminar con el último intervalo.

En Excel es mucho más sencillo de hacer, para lo cual recurrimos a la función *Frecuencia*, que se explica a detalle en un archivo anexo, o bien mediante la ayuda de esta función directamente en la hoja de cálculo.

#### 4.5. Frecuencia Relativa ( $F_i \%$ ):

Esta columna nos dice el porcentaje con que contribuye cada intervalo en la distribución total. La frecuencia relativa se encuentra dividiendo la frecuencia absoluta de cada intervalo, entre el número total de datos.

$$\Sigma FR = 100\%$$

#### 4.6. Frecuencia Acumulada “menos qué”:

Esta frecuencia nos ayuda a cuantificar los datos que se van *acumulando* intervalo tras intervalo.

La frecuencia acumulada del primer intervalo, siempre será igual a la frecuencia absoluta del mismo, ya que

solamente se han acumulado esos datos. A partir del posterior (intervalo 2), se debe sumar, a la frecuencia absoluta de ese intervalo, la frecuencia acumulada del intervalo anterior; en este caso, del primero.

Tras seguir el mismo procedimiento hasta culminar con todos los intervalos, observamos que la última frecuencia acumulada es igual al número total de datos de la distribución, pues evidentemente, se han sumado todos y cada uno de ellos de forma acumulada.

Para comprenderlo de mejor forma, observemos el siguiente ejemplo:

Nº Intervalo	Límites de Clase		Límites Reales de Clase		Marca de Clase	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Acumulada	Frecuencia Acumulada %
	L.I.C.	L.S.C.	L.R.I.C.	L.R.S.C.	$X_i$	$f_i$	$f_i \%$	"- que "	"- que %"
1	2.86	4.06	2.855	4.065	3.46	11	26.83%	11	
2	4.07	5.27	4.065	5.275	4.67	14	34.15%	25	
3	5.28	6.48	5.275	6.485	5.88	9	21.95%	34	
4	6.49	7.69	6.485	7.695	7.09	7	17.07%	41	
						<b>41</b>	<b>100%</b>		

✚ Como podemos apreciar, la *Frecuencia Acumulada* en el primer intervalo, corresponde a la *Frecuencia Absoluta*, lo cual implica que hubo 11 datos cuyo valor oscila entre 2.86 y 4.06.

✚ Para el segundo intervalo, tenemos que se han **acumulado** otros 14 datos, por lo tanto, la *Frecuencia Acumulada* se incrementa a 25.

Nº Intervalo	Límites de Clase		Límites Reales de Clase		Marca de Clase	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Acumulada	Frecuencia Acumulada %
	L.I.C.	L.S.C.	L.R.I.C.	L.R.S.C.	$X_i$	$f_i$	$f_i \%$	"- que "	"- que %"
1	2.86	4.06	2.855	4.065	3.46	11	26.83%	11	
2	4.07	5.27	4.065	5.275	4.67	14	34.15%	25	
3	5.28	6.48	5.275	6.485	5.88	9	21.95%	34	
4	6.49	7.69	6.485	7.695	7.09	7	17.07%	41	
						<b>41</b>	<b>100%</b>		

✚ Ahora bien, podemos apreciar que esos mismos 25 datos que se han acumulado hasta el segundo intervalo, tienen un valor que es **menor que 5.27** (Ver *L.S.C. del intervalo 2*). Dicho de otra forma, existen 25 datos cuyo valor oscila entre 2.86 (L.I.C. del primer intervalo) y 5.27.

✚ Como es lógico de pensar, la *Frecuencia Acumulada* del último intervalo, nos da como total 41 datos, o sea el número total de observaciones registradas. Esto implica que existen 41 datos cuyo valor es menor que 7.69 (L.S.C. del último intervalo). En otras palabras, existen 41 datos cuyo valor oscila entre 2.86 y 7.69.

Como podemos apreciar, para poder interpretar esta frecuencia, es indispensable vincularla con el Límite Superior de Clase del intervalo en cuestión.

#### 4.7. Frecuencia Acumulada Relativa “menos qué %”:

Esta columna nos indica el porcentaje que representa la frecuencia acumulada de cada intervalo con respecto a la distribución total.

La frecuencia acumulada relativa se encuentra dividiendo la frecuencia acumulada de cada intervalo, entre el número total de datos.

#### 4.8. Frecuencia Complementaria “más qué”:

La Frecuencia Complementaria es la inversa de la Acumulada; es decir, partiendo del total de datos, que sería el valor que asume el primer intervalo, nos va indicando cómo se van segregando los datos (disminuyendo) conforme se va avanzando en los intervalos. Nos sirve para saber cuanto falta aún de la distribución.

Se utiliza la notación “más que” ó “más de” y hace intervenir el límite inferior de clase.

Se calcula de la siguiente forma: la Frecuencia Complementaria del primer intervalo siempre será igual a “n”; es decir, el número total de datos observados. Para los siguientes intervalos, se deberá restar a la frecuencia complementaria del intervalo anterior, la frecuencia absoluta del intervalo en cuestión.

Tras seguir el mismo procedimiento hasta culminar con todos los intervalos, observamos que la última frecuencia complementaria es igual a la frecuencia absoluta del último intervalo, pues evidentemente, son los datos que quedarían por eliminar de la distribución para llegar a cero.

Veamos el siguiente ejemplo ilustrativo:

Nº Intervalo	Límites de Clase		Límites Reales de Clase		Marca de Clase	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Complementaria	Frecuencia Complementaria %
	L.I.C.	L.S.C.	L.R.I.C.	L.R.S.C.	$X_i$	$f_i$	$f_i \%$	" + que "	" + que "
1	2.86	4.06	2.855	4.065	3.46	11	26.83%	41	100%
2	4.07	5.27	4.065	5.275	4.67	14	34.15%	30	73%
3	5.28	6.48	5.275	6.485	5.88	9	21.95%	16	39%
4	6.49	7.69	6.485	7.695	7.09	7	17.07%	7	17%
						<b>41</b>	<b>100%</b>		

- ✚ El número total de datos es 41, o sea, la primer frecuencia complementaria.
- ✚ Para obtener la subsecuente, restamos a este valor (41), la frecuencia absoluta del intervalo anterior (11), por lo que la frecuencia complementaria del intervalo 2, resulta ser 30.
- ✚ Esto implica que el **complemento** de nuestra distribución, son los restantes 30 datos, del total de 41 que teníamos.



Nº Intervalo	Límites de Clase		Límites Reales de Clase		Marca de Clase	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Complementaria	Frecuencia Complementaria %
	L.I.C.	L.S.C.	L.R.I.C.	L.R.S.C.	$X_i$	$f_i$	$f_i \%$	" + que "	" + que "
1	2.86	4.06	2.855	4.065	3.46	11	26.83%	41	100%
2	4.07	5.27	4.065	5.275	4.67	14	34.15%	30	73%
3	5.28	6.48	5.275	6.485	5.88	9	21.95%	16	39%
4	6.49	7.69	6.485	7.695	7.09	7	17.07%	7	17%
						<b>41</b>	<b>100%</b>		

- Como podemos apreciar, tenemos 30 datos cuyo valor es *mayor a* 4.07 (L.I.C. del intervalo 2). En otras palabras, existen 30 datos cuyo valor oscila entre *más de* 4.07 y 7.69.

#### 4.9. Frecuencia Complementaria Relativa "más qué %":

Al igual que en el resto de las frecuencias relativas, esta columna nos indica el porcentaje que representa la frecuencia complementaria de cada intervalo con respecto a la distribución total.

La frecuencia complementaria relativa se encuentra dividiendo la frecuencia acumulada de cada intervalo, entre el número total de datos (Ver tabla anterior).