



San Fernando College TP  
Asignatura: Matemática  
Prof. Franco Cabezas Castro

$v_o B_o$  UTP

# Probabilidad y Estadística: Guía II

Nombre:

Curso:

Fecha:

**Objetivo: Describir y representar datos, tablas y gráficos**

## 1. Introducción

En esta guía vas a aprender cómo recopilar información para identificar y analizar un problema. Vas a aprender también cómo organizar esta información para poder interpretarla. Asimismo, de qué manera puedes presentarla gráficamente una vez que la has organizado.

Por último, vas a saber qué es un histograma, qué es un polígono de frecuencias y, lo más importante, que más allá de ser representaciones visuales de una masa de datos, tienen propiedades estadísticas que te proporcionan mucha información valiosa.

Estos conocimientos te permitirán entrar de lleno en el mundo de la Estadística.

## 2. Organización de los datos

Existen muchas formas de organizar los datos. Para poder decidir cuál es la mejor debes tener muy claro para qué quieres la información. Recuerda a Bachelard:

Lo importante no es medir para pensar, sino pensar para medir.

### 2.1. ¿Para qué te sirve tener organizada la información?

Imagínate que tu profesor te pide que hagas una presentación de tres minutos frente a tu grupo sobre tu comunidad.

¿Crees que puedes hablar de todo lo que hay en tu comunidad? Si no, ¿de qué hablarías? ¿Hablarías de su tamaño, de sus habitantes, de su vegetación y clima, de sus actividades productivas, de lo que haces como miembro de tu comunidad o hablarías de tu colegio, de lo que más te gusta, de lo que te gusta menos?

Toda esta información sobre tu comunidad es valiosa, pero ¿a qué características les darías prioridad al hablar? ¿Por qué? ¿Cómo la organizarías?

Supón que decides trabajar sobre las edades de los habitantes de tu comunidad y que su población es de 2.500 personas. Como no tienes el tiempo de ir a preguntarle su edad a cada uno de ellos, decides salir a la calle y entrevistar a las primeras 50 personas que te encuentras en el camino.

Los datos que encontraste son los siguientes:

Edades de los 50 encuestados en la comunidad				
2	17	6	5	14
6	43	3	26	26
31	9	24	12	8
33	13	6	39	32
23	32	30	6	30
18	11	22	36	21
27	11	16	24	13
13	14	35	23	20
37	28	21	22	29
21	20	29	13	23

## Actividad de aprendizaje I

¿Qué observas en este conjunto de datos? Describe al menos cinco características que encuentres.

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_

### 2.2. Formas de organización de datos

Tomemos los datos de edades de la sección anterior. ¿Cómo ordenarías la información?

No existe una forma única de ordenar los datos. Puede ser de menor a mayor, de mayor a menor, por decenas, entre otras. ¿Cuál es la mejor? La que te convenga, de acuerdo con el propósito de tu investigación.

## Actividad de aprendizaje II

Selecciona una de las formas de ordenar la información de edades en la siguiente tabla, la cual puede ser de: mayor a menor, menor a mayor, por grupos de edades repetidas (por columna). Argumenta por qué tu tabla permite conocer de manera más directa:

- ¿Cuántas personas están en el rango de edad de 20 a 30 años?
- ¿Cuál es el mayor grupo de personas con la misma edad?
- ¿Encuentras de forma más fácil las cinco características que describiste en la actividad anterior?

Edades ordenadas de los 50 encuestados en la comunidad				

### 2.3. Tipos de frecuencia: absoluta, relativa, acumulada y relativa acumulada

Si te fijas únicamente en los resultados para las edades de 24, 26, 27 y 28 años observarás lo siguiente:

Edad	Número de casos	Total
24	1	2
26	1	2
27	1	1
28	1	1

A este cuadro se le llama **cuadro de conteo**.

El cuadro de conteo te permite resumir la nube de datos. Ahora que sabes esto, elabora el cuadro de conteo de todos los datos de edades de las personas que entrevistaste.

Cuadro de conteo de las edades de 50 habitantes de mi comunidad		
Edad	Número de casos	Total ( $f_a$ )



El cuadro de conteo te permite obtener información para organizar y facilitar el conteo de los datos. Al total de observaciones de cada renglón (edad en nuestro ejemplo) de tu cuadro de conteo se le llama frecuencia absoluta y se denota como  $f_a$ . En tu cuadro puedes observar, por ejemplo, que la frecuencia absoluta para la edad de 26 es de dos, mientras que para la de 27 es solo de uno. Como puedes ver, la columna de total del cuadro de conteo se llama **fat (frecuencia acumulada total)**.

### Actividad de aprendizaje III

A partir de tu cuadro, identifica los siguientes grupos de personas:

- La población que está cursando algún grado de educación básica, considerando que, en condiciones óptimas, inician a los 6 años y concluyen a los 14.
- La población que está cursando algún grado de educación media, considerando que, en condiciones óptimas, inician a los 15 años y concluyen a los 18.
- La población en edad de votar.

1. ¿Cuántas personas hay en cada grupo?

Número de habitantes que:		
Cursan la primaria	Cursan la secundaria	Están en edad de votar

2. ¿Consideras que obtuviste del cuadro de conteo información más precisa de la que ya tenías sobre tu comunidad y sus habitantes? Argumenta tu respuesta.

---

---

3. Ahora supón que tu profesor te solicita que le presentes un breve informe al alcalde sobre el número de habitantes que pudieran recibir un apoyo económico del programa “Crece más”. Este programa del gobierno del estado atiende a las personas menores de 17 años. ¿Cómo obtendrías esta información?

---

---

4. ¿En qué fue diferente la forma en que obtuviste la información sobre el programa “Crece más” y sobre la población en edad de votar?

---

---

### 3. Análisis de datos a través de tablas

Seguramente te habrás dado cuenta de que para resolver el primer caso tuviste que sumar las frecuencias de la población de entre 2 y 16 años (los menores de 17 años). Para el segundo caso, las frecuencias de los que tienen 18 años o más. Sabiéndolo o no, en ambos casos utilizaste un nuevo concepto: **el de frecuencia absoluta acumulada ( $f_{aa}$ )**

La frecuencia absoluta acumulada se obtiene de sumar la frecuencia absoluta de los datos anteriores con la faa del dato actual.

#### Actividad de aprendizaje IV

Como puedes observar en el cuadro siguiente, agregamos una nueva columna, la de frecuencia absoluta acumulada ( $f_{aa}$ ).

Llena la columna de frecuencia absoluta acumulada, según como se indica en la tabla:

Edades de 50 habitantes de mi comunidad		
Edad	Frecuencia absoluta ( $f_a$ )	Frecuencia absoluta acumulada ( $f_{aa}$ )
2		
3		
5		
6		
8		
9		
11		
12		
13		



### Edades de 50 habitantes de mi comunidad

Edad	Frecuencia absoluta ( $f_j$ )	Frecuencia absoluta acumulada ( $f_{j\alpha}$ )	$f_{j\alpha}$ en personas de 18 años o más
14			
16			
17			
18			
20			
21			
22			
23			
24			
26			
27			
28			
29			
30			
31			
32			
33			
35			
36			
37			
39			
43			
<b>Total (fat)</b>		9	

En el caso de la población en edad de votar, si hubieras tomado la frecuencia acumulada total y le hubieras restado la frecuencia acumulada hasta antes de la edad de 18 años tendrías el mismo resultado que obtuviste llenando la cuarta columna del cuadro.

En estos dos ejercicios resolvimos dos preguntas: en el caso del programa “Crece más” contestamos a la pregunta **menos de**, directamente con la tabla de frecuencias acumuladas. En el segundo caso, el de la población en edad de votar, resolvimos la pregunta **mayor de** con la fórmula:

$$\text{edad mayor de} = \text{edad total} - \text{observaciones acumuladas hasta antes del dato que buscamos}$$

El informe que presentaste al alcalde le gustó. Por ello te ha pedido que realices el mismo ejercicio para los habitantes de la comunidad “San Fernandito”. Te pide que compares los resultados que obtengas con los de tu comunidad.

Con la información que le proporciones piensa establecer el monto de recursos que van a utilizar en el programa “Crece más” en esa localidad. Como en el caso anterior te asignó un monto para que realices la investigación. Sin embargo, otra vez no te alcanzan ni el tiempo ni los recursos para entrevistar a todos los habitantes de “San Fernandito”. Como la población de esta comunidad es cinco veces mayor que la tuya, decidiste encuestar cinco veces más personas que en el caso de tu comunidad, es decir 250 personas. Para hacer tu investigación tomaste el bus hasta la comunidad de “San Fernandito”, te bajaste en el terminal y entrevistaste a las primeras 250 personas que te encontraste.

Los datos que obtuviste en la encuesta son los siguientes:

Edades de 250 habitantes de "El Capulín"																								
51	7	89	74	53	69	22	83	7	11	75	83	17	24	60	61	87	84	63	8	7	52	31	48	84
74	70	3	37	35	12	52	29	30	49	3	75	19	5	77	63	80	37	37	69	90	48	50	45	10
1	83	87	20	55	5	63	59	38	27	84	8	79	31	36	61	833	38	88	26	79	17	51	86	9
72	82	30	88	69	19	43	52	44	76	48	73	12	76	85	7	28	5	48	6	68	13	79	81	89
8	8	17	78	13	43	76	86	4	28	84	55	81	40	11	78	83	45	32	67	50	18	78	2	88
85	65	11	63	14	81	86	76	3	25	44	21	19	68	70	41	28	63	65	84	14	45	80	44	76
68	37	65	16	33	20	10	81	32	38	76	76	49	71	53	6	32	17	19	86	8	43	77	27	7
73	61	62	67	4	44	28	11	88	88	9	64	90	56	10	21	34	53	37	89	37	5	54	73	84
48	71	15	33	35	50	56	71	21	21	9	90	36	60	72	43	47	16	50	17	42	58	57	81	44
80	51	27	66	50	89	13	62	44	44	64	19	50	5	34	84	88	23	4	60	26	20	77	53	86



2. ¿Qué semejanzas y diferencias encuentras al comparar ambas poblaciones: la de tu comunidad y la de “San Fernandito”?

---

---

---

3. ¿Consideras que esta muestra de 250 personas tiene el mismo valor que el de las edades de las 50 personas que realizaste en tu comunidad? Argumenta tu respuesta.

---

---

---

4. ¿Cómo puedes comparar los resultados de dos comunidades de tamaños diferentes? ¿Tiene más valor hacer muestras más grandes?

---

---

---

La comparación de resultados en las dos comunidades de tamaño diferente la puedes hacer obteniendo para ambas los porcentajes de frecuencia de cada una de las edades registradas. Hacer esto se llama obtener las *frecuencias relativas*.