



BIENVENIDOS A SU CUARTO  
CUATRIMESTRE ESTIMADOS  
LICENCIADOS EN ENFERMERÍA

MATERIA: BIOESTADISTICA

ORIENTADOR: ROSARIO GÓMEZ LUJANO



## PRIMER MODULO

Del 10 de septiembre al 22 de octubre del año 2022

Evaluación del parcial Sábado 22 de octubre del año 2022

### Criterios de evaluación

Actividad 1: 25%

Fecha limite para enviar la actividad 26/09/2022

Actividad 2: 25%

Fecha limite para enviar la actividad 17/10/2022

Evaluación: 50%



**UNIDAD I: ESTADISTICA DESCRIPTIVA**

**UNIDAD II: CALCULO DE PROBABILIDADES**

**UNIDAD III: DISTRIBUCIONES DE PROBABILIDAD**

**UNIDAD IV: DEMOGRAFIA**

**OBJETIVO DE LA ASIGNATURA:** Lograr que el estudiante conozca y aprenda a utilizar las técnicas estadísticas y sus aplicaciones más frecuentes en las distintas áreas de Salud.

¿Qué es la estadística?

¿Qué es bioestadística?

¿Cómo se clasifica la  
estadística?

# Conceptos básicos de estadística descriptiva

La **estadística** es la ciencia que trata de la recolección, organización, presentación, análisis e interpretación de datos numéricos con el fin de realizar una forma de decisión más efectiva. Para su mejor estudio, se ha dividido en dos ramas las cuales son: estadística descriptiva y estadística inferencial.

## ESTADÍSTICA DESCRIPTIVA

Se ocupa de la **recopilación, organización y análisis y presentación de datos** sobre alguna característica de ciertos individuos pertenecientes a una población. La información es presentada en forma de tablas y gráficas.

## ESTADÍSTICA INFERENCIAL

Se **encarga de hacer deducciones**, es decir, inferir propiedades, conclusiones y tendencias, a partir de una muestra del conjunto. Su papel es interpretar, hacer proyecciones y comparaciones.



## En la estadística descriptiva los conceptos relevantes son:

**La bioestadística** es una rama de la estadística que se ocupa de los problemas planteados dentro de las ciencias de la vida, como la biología, la medicina, la enfermería, entre otras.

**Población:** Conjunto bien definido sobre el que se observa cierta característica. Este conjunto puede ser finito o infinito. Así, el tamaño de la población es el número de individuos que tiene, denotado por  $N$ .

**Individuo:** Es cada uno de los elementos de la población.

**Muestra:** Cuando la población es muy grande, se recurre a una muestra, que es un subconjunto de individuos de una población, que refleja las características de esta lo mejor posible. El tamaño de la muestra lo denotamos por  $n$ .

**Variable:** Es una característica que se desea estudiar de una población. Por ejemplo: el deporte favorito de los alumnos de primer semestre, la estatura de las mujeres.

Una  
**Variable**  
puede ser

Cualitativa o  
de atributos

◆ Nominal

◆ Ordinal

Cuantitativa  
o numérica

◆ Discreta

◆ Continua

Así tenemos que:

- Las variables numéricas **discretas** son aquellas que obedecen a un conteo y toman valores enteros.
- Las variables numéricas **continuas** son las que obedecen a mediciones y toman valores continuos.
- Las variables categóricas **nominales** son las que toman valores que se corresponden con cualidades no cuantificables de los individuos.
- Las variables categóricas **ordinales** son las que presentan una relación de orden entre sus valores.

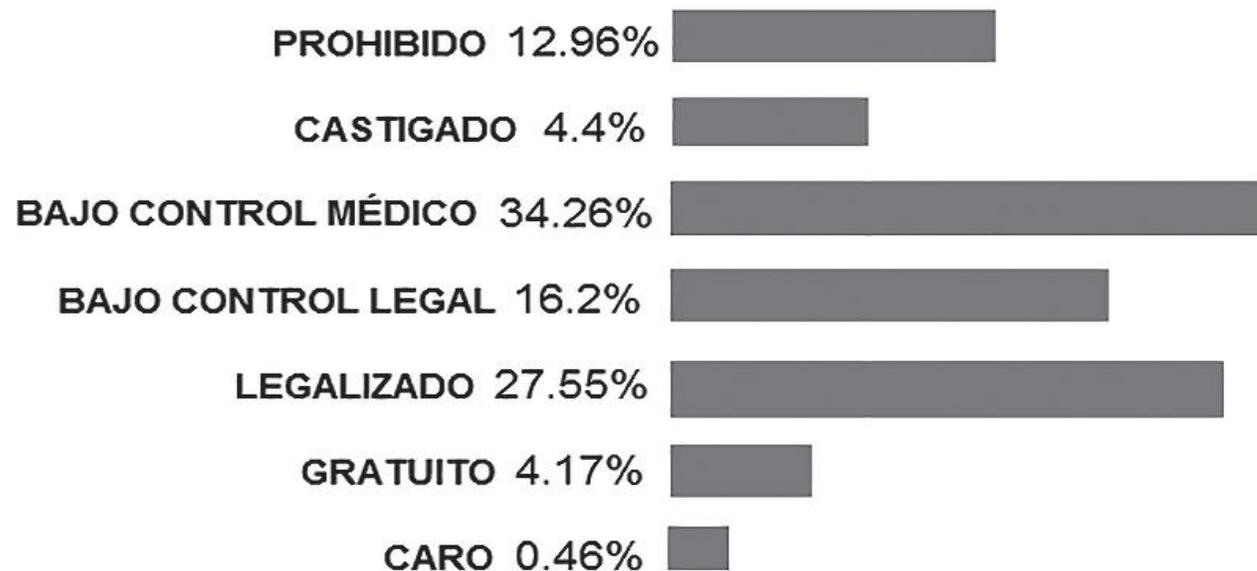
## Información estadística publicada en revistas

### SOCIEDAD

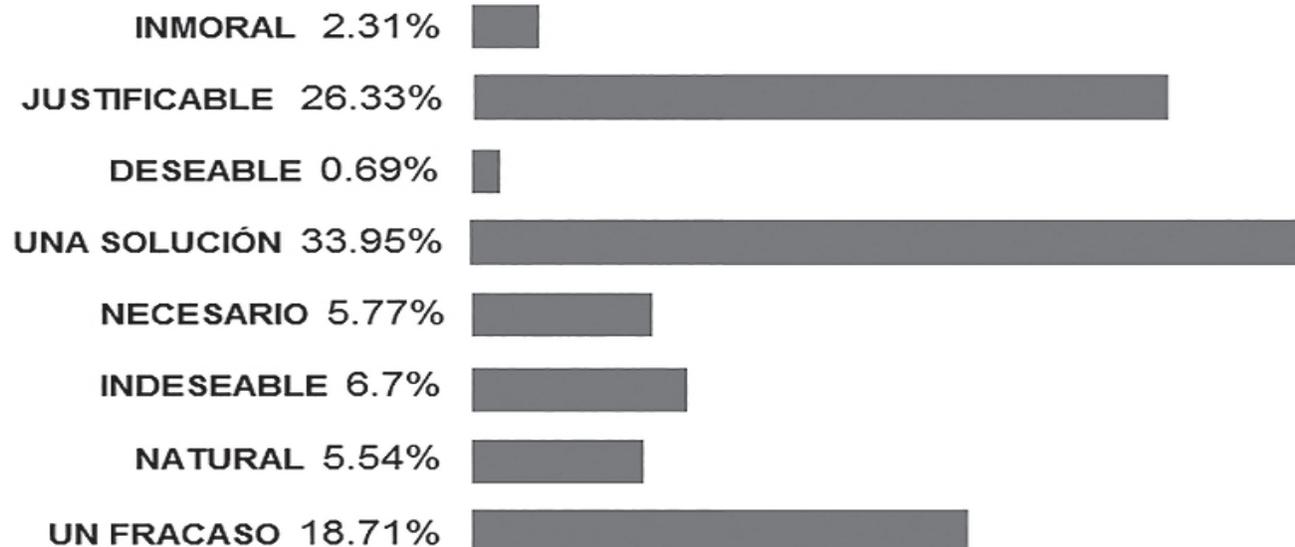
#### *¿Conservadores o liberales?*

**Día Siete** realizó una consulta, a través de su página de Internet, para conocer las opiniones de los lectores respecto a temas como aborto, divorcio y características que debe tener una persona. Cerca de 500 personas participaron.

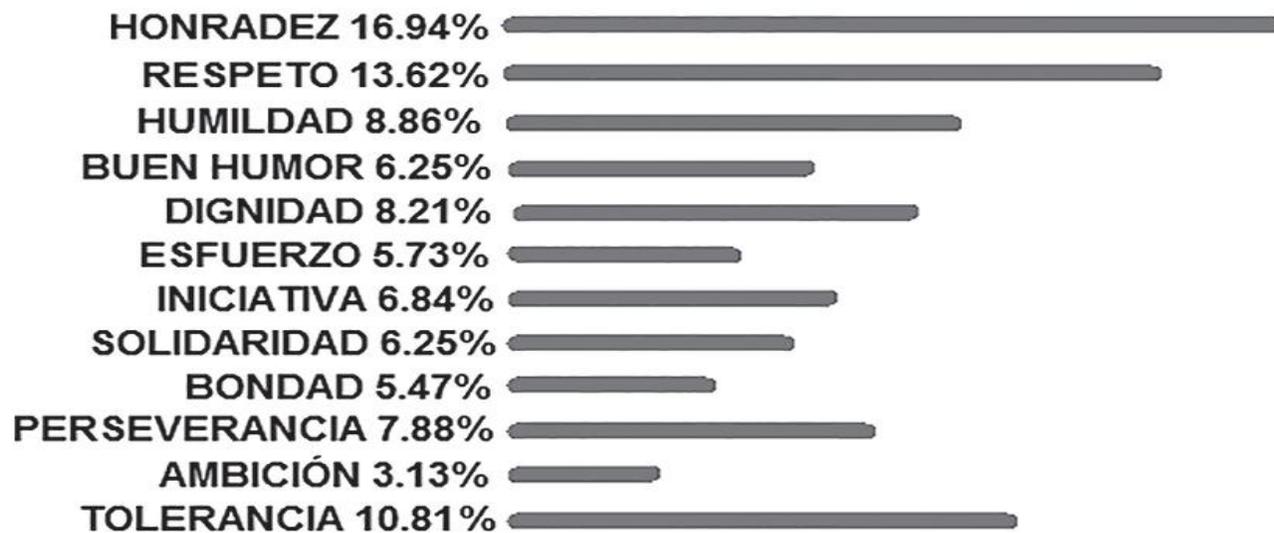
### EL ABORTO DEBE SER...



EL DIVORCIO ES...



¿QUÉ CARACTERÍSTICA  
CONSIDERAS MÁS VALIOSA  
EN UNA PERSONA?



## ¿Estadística descriptiva o inferencial?

1. Al observar el funcionamiento de cuatro pilas alcalinas AA marca 'solar' se sabe que éstas dejaron de funcionar después de 5, 6.5, 7 y 4.5 horas de uso continuo, también que cuatro pilas alcalinas AA marca 'inergy' dejaron de funcionar después de 5, 7.8, 7 y 6.2 horas de uso continuo respectivamente. ¿Cuáles de las siguientes conclusiones provienen de la estadística descriptiva y cuáles de la estadística inferencial? Elige la respuesta correcta y justifica tu elección.

**A) La diferencia entre los tiempos promedios de funcionamiento de las dos marcas es 0.75**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN:

**B) Si se selecciona y prueba otra pila alcalina AA marca 'inergy' seguramente durará más que una de marca 'solar'.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACION

**C) El funcionamiento promedio de las cuatro pilas alcalinas AA 'solar' fue de 5.75 horas.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN:

**D) Las pilas alcalinas AA 'inergy' funcionan más tiempo que las pilas alcalinas AA 'solar'.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN:

2. En tres días consecutivos, un policía de tránsito infraccionó a 6, 10 y 12 conductores por no respetar un señalamiento de ALTO y a 15, 8 y 18 personas por conducir en exceso de velocidad. ¿Cuáles de las siguientes conclusiones provienen de la estadística descriptiva y cuáles de la estadística inferencial? Elige la respuesta correcta y justifica tu elección.

**A) En los tres días el policía infraccionó a más personas por conducir con exceso de velocidad que por no respetar un señalamiento de ALTO.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN:

**B) Si se infracciona próximamente a un conductor será por conducir en exceso de velocidad antes que por no respetar un señalamiento de ALTO.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACION

**C) El primer día, se levantaron 9 infracciones más por conducir en exceso de velocidad que por no respetar un señalamiento de ALTO.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN:

**D) Es necesario colocar más avisos de “VELOCIDAD MÁXIMA” en las carreteras.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN

3. Los registros del departamento de bomberos de una población rural muestran que durante los meses de enero a diciembre del año pasado sofocaron 5, 4, 2, 6, 8, 12, 18, 15, 6, 8, 4 y 2 incendios forestales respectivamente. De las siguientes conclusiones ¿cuáles describen únicamente y cuáles implican una generalización o inferencia?

**A) El año pasado se sofocaron en promedio 7.5 incendios forestales cada mes.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN

**B) Si ocurre un incendio seguramente será durante los meses de verano.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACION

**C) El departamento de bomberos necesitará más voluntarios de junio a agosto que el resto del año.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN:

**D) En el mes de enero ocurrieron 3 incendios más que en el mes de diciembre.**

Conclusión Descriptiva

Conclusión Inferencial

JUSTIFICACIÓN:

**Analiza e identifica las variables que se piden según su clasificación:**

1. Una muestra de 100 clientes de un salón de belleza fue cuestionada en cuanto a su color favorito de cabello, la colonia en la que habitan y el nivel de satisfacción respecto a la atención que recibieron en el salón de belleza. Identifica las variables (si es que existen):

VARIABLE NOMINAL

VARIABLE DISCRETA

VARIABLE ORDINAL

VARIABLE CONTINUA

2. La velocidad máxima es de 40 km/hora en una zona residencial y al menos 300 autos transitan por ahí diariamente; los residentes se sienten angustiados por el exceso de tráfico y han decidido colocar al menos 10 avisos de 'maneje con precaución' en la zona. Identifica las variables (si es que existen):

VARIABLE NOMINAL

VARIABLE DISCRETA

VARIABLE ORDINAL

VARIABLE CONTINUA



**OBSERVACIÓN**, técnica que se basa en la anotación y registro de acciones, reacciones y en general del comportamiento observado de los sujetos estudiados. El ambiente de los individuos, no se altera o modifica mientras se lleva a cabo la observación.

**ENCUESTA**, técnica que se basa en la captación directa de las opiniones y puntos de vista de los individuos sujetos de estudio, mediante el registro de sus respuestas, a lo que el encuestador se limita. El *questionario* es el instrumento que se utiliza en la encuesta.

**EXPERIMENTACIÓN**, consiste en 'provocar' las situaciones o eventos que darán pie al análisis del comportamiento de los sujetos. Al contrario de la observación, con esta técnica se recrea el ambiente en que los individuos se desenvuelven.

**ENTREVISTA**, comprenden la interacción verbal directa ente los individuos sujetos de estudio y el entrevistador. El entrevistado responde a un cuestionario guía pero profundiza a medida que la charla transcurre.

- **Guía de observación:** se emplea básicamente para recabar datos del comportamiento o conducta de un sujeto, ya sea de forma individual o en grupo, en situaciones reales.
- **Escala de estimación:** este instrumento de evaluación pretende identificar la frecuencia o intensidad en la que se presenta una conducta, o los niveles de aceptación de un hecho o fenómeno, mediante una escala que puede ser numérica, descriptiva o categórica.
- **Registro descriptivo:** se emplea para obtener información sobre el comportamiento del individuo, recabando evidencias de su actividad, sea esta positiva o negativa en un contexto en el que previamente esté definida la competencia en cuestión.
- **Lista de cotejo:** consiste en un listado de aspectos a evaluar, ya sean contenidos, capacidades, habilidades o conductas, que habrán de calificar si se lograron o no.
- **Cuestionario:** instrumento de indagación con una serie de cuestionamientos de base estructurada o abierta. Permite obtener, de forma rápida, información de un sujeto o de una población determinada; permite hacer comparaciones entre grupos e individuos.





Una Distribución o Tabla de Frecuencias es la representación conjunta de los datos en forma de tabla o subgrupo de datos correspondientes a un fenómeno en estudio, y su ordenamiento en base al número de observaciones que corresponden a cada dato o a cada grupo de datos, adecuados según cronología, geografía, análisis cuantitativo o cualitativo. La presentación de los datos en forma ordenada, por medio de una tabla, dependerá de los datos de que se trate, y si estos son cualitativos o cuantitativos como se muestra a continuación

### **Cualitativos**

Alfabético A – Z

Alfabético Z – A

Del más al menos repetido

Del menos al más repetido

### **Cuantitativos**

Creciente (menor al mayor)

Decreciente (mayor al menor)



Se preguntó a un grupo de alumnos de primer año del Cobach Pichucalco, por la asignatura de su preferencia, arrojándose los siguientes resultados:

## Asignaturas

Mate Social Taller Quím. Infor Mate Inglés Mate Quím. Infor Inglés Ética Inglés Social Inglés Ética Mate Taller Quím. Mate Taller Social Mate Inglés Infor Inglés Ética Infor Mate Inglés Infor Ética Quím. Taller Inglés Social Inglés Ética Taller Infor Quím. Taller Taller Infor Mate Quím. Infor Mate Infor Inglés

DATOS CUALITATIVOS

Realizar una distribución de frecuencia.

Se aplicó un examen de matemáticas a un grupo de 25 alumnos de primer año del Cobach Pichucalco, y se obtuvieron las siguientes calificaciones.

5,6,5,7,8,9,10,5,7,8,7,8,9,10,5,6,,6,8,  
9,10, 5,6,7,8,9 datos cuantitativos



## Frecuencia absoluta, absoluta acumulada, relativa y relativa acumulada.

**Frecuencia Absoluta de un dato** es el número de veces que se repite ese dato, también se presenta la frecuencia absoluta de un intervalo que se refiere al número de datos que pertenecen a ese intervalo. La denotaremos por  $f$ .

**Frecuencia Absoluta Acumulada:** Hasta un dato específico, es la suma de las frecuencias absolutas de todos los datos anteriores, incluyendo también la del dato mismo del cual se desea su frecuencia acumulada. De un intervalo es la suma de las frecuencias absolutas de todos los intervalos de clase anteriores, incluyendo la frecuencia del intervalo mismo del cual se desea su frecuencia acumulada. La denotaremos por  $a_{fa}$ . La última frecuencia absoluta acumulada deberá ser igual al número total de datos.

**Frecuencia Relativa:** De un dato, se obtiene al dividir la frecuencia absoluta de cada dato entre el número total de datos. De un intervalo se obtiene al dividir la frecuencia absoluta de cada intervalo entre el número total de datos. La denotamos por  $fr$ .

**Frecuencia Relativa Acumulada:** Hasta un dato específico, es la suma de las frecuencias relativas de todos los datos anteriores, incluyendo también la del dato mismo del cual se desea su frecuencia relativa acumulada. De un intervalo es la suma de las frecuencias relativas de todos los intervalos de clase anteriores incluyendo la frecuencia del intervalo mismo del cual se desea su frecuencia relativa acumulada, La denotaremos por  $fra$ . La última frecuencia relativa acumulada deberá ser igual a la unidad.

## Examen de CI

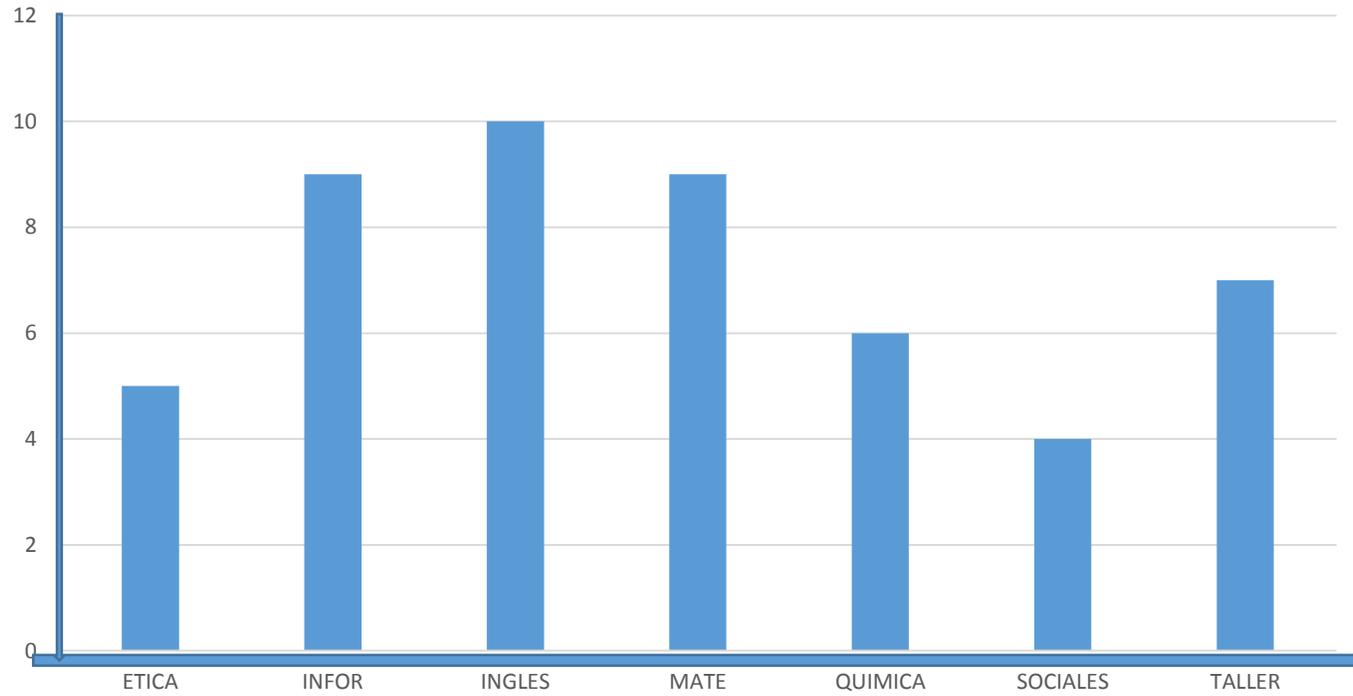
MATERIA	F	Fa	Fr	Fra
ETICA	5	5	5/50=0.1	0.1
INFOR	9	14	0.18	0.28
INGLES	10	24	0.2	0.48
MATE	9	33	0.18	0.66
QUIMICA	6	39	0.12	0.78
SOCIALES	4	43	0.08	0.86
TALLER	7	50	0.14	1
<b>TOTAL</b>	<b>50</b>		1	

$$50=100\%$$

$$9=$$

$$\frac{9 \times 100}{50} = 18\%$$

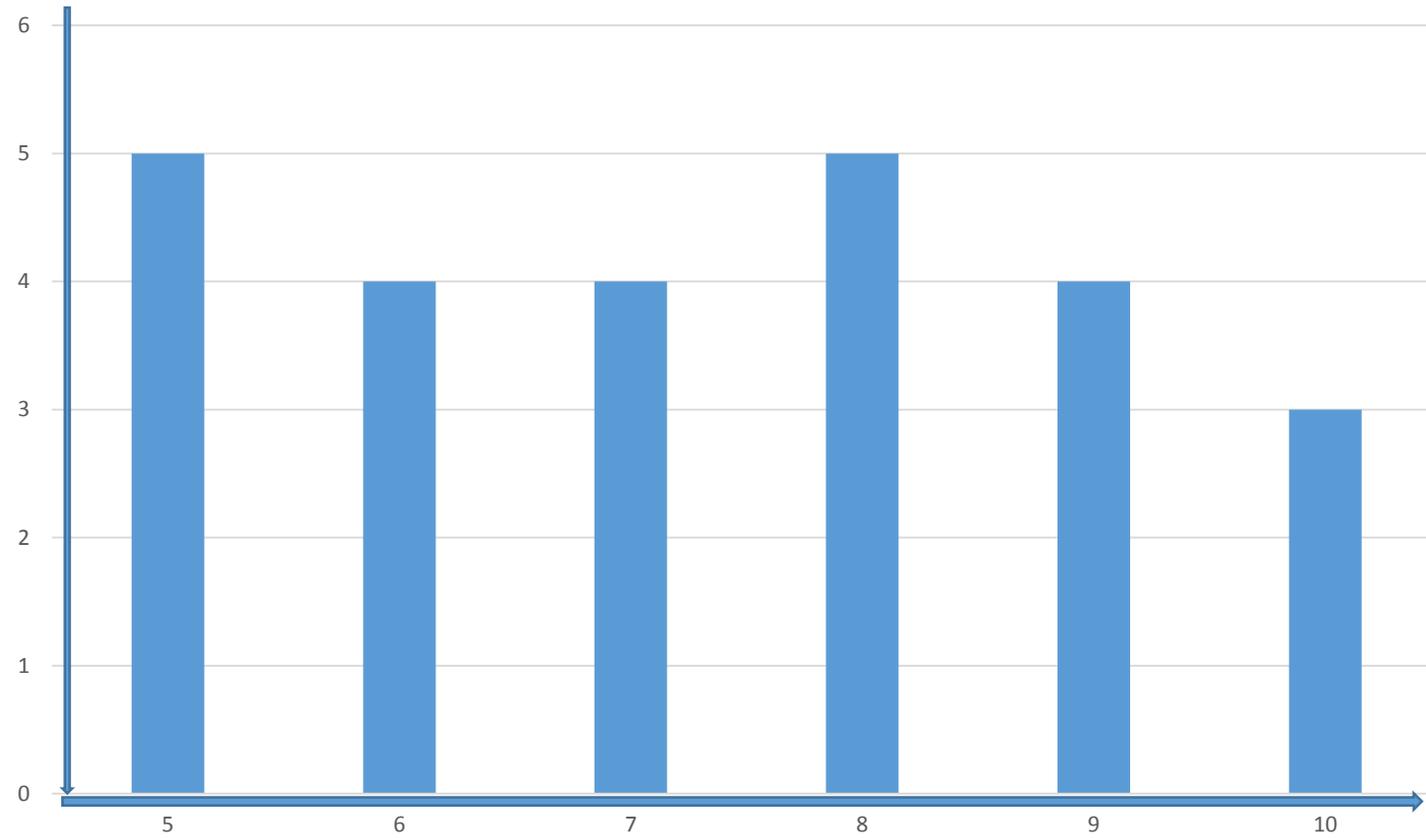
Materias preferidas



## Calificaciones de matemáticas

Cal	F	Fa	Fr	Fra
5	5	5	0.2	0.2
6	4	9	0.16	0.36
7	4	13	0.16	0.52
8	5	18	0.2	0.72
9	4	22	0.16	0.88
10	3	25	0.12	1
TOTAL	25		1	

Calificaciones de matematicas



Se han investigado los precios de las habitaciones de 40 hoteles de cierta ciudad y hemos recogido los resultados que a continuación detallamos:

200	300	500	600	400	700	1000	300
400	700	1000	200	200	600	600	300
700	1000	700	600	300	400	1000	700
300	500	700	1000	200	400	600	500
500	500	400	300	600	700	1000	200

Construye una tabla de distribución de frecuencia.

1.-¿Qué porcentaje de hoteles cobra 700 pesos por noche?

2.-¿Qué porcentaje de hoteles cobra 200 pesos por noche?

## Medidas de Centralización

- a) Media Aritmética o promedio  $\bar{x}$
- b) Mediana  $\tilde{x}$
- c) Moda  $\hat{x}$

## Medidas de Variabilidad o Dispersión

- a) Rango ( R )
- c) Varianza ( $s^2$ )
- d) Desviación estándar o típica ( S )

## Medidas de tendencia central para datos no agrupados

Llamaremos datos no agrupados a los que no aparecen resumidos en distribuciones de frecuencias.

a) **Media Aritmética.** La medida más evidente que podemos calcular para describir un conjunto de observaciones numéricas es su valor medio. La media no es más que la suma de todos los valores de una variable dividida entre el número total de datos de los que se dispone. Siendo su fórmula la siguiente.

$$\text{MEDIA ARITMETICA} = \frac{\sum x}{n}$$

**Donde:**

**$\Sigma$**  Símbolo de sumatoria que indica que se deberá sumar todos los valores que toma la variable numérica X.

**X** Cada uno de los datos obtenidos de la muestra.

**n** Número total de datos

## **b) Mediana**

Otra medida de tendencia central o de centralización que se utiliza habitualmente es la mediana. Es el dato o valor equidistante o que se encuentran más en medio de todo el conjunto de datos numéricos.

Para obtener la mediana para datos no agrupados primeramente debemos ordenar los datos en forma ascendente o descendente observando el dato que esta exactamente a la mitad. En dado caso que encontremos dos números que dividen al conjunto sumaremos los dos números y lo dividimos entre dos.

# Moda

La moda es una medida de tendencia central que describe “lo más común” o “lo que más se repite” (lo más frecuente) en un conjunto de datos que pueden ser numéricos o cualitativos. La moda puede no existir (si ningún dato se repite) y si existe, entonces puede ser unimodal, bimodal o multimodal.

## Medidas de variabilidad

Conocidas también como medidas de dispersión o descentralización, estas medidas o puntos de referencia obtenidos, se utilizan para analizar los valores numéricos de la muestras y saber qué tan dispersos están con respecto a su media, entre las más comunes están ***el rango, varianza y desviación estándar.***

### Rango o recorrido

Es una medida de dispersión que solamente indica “la distancia” que existe entre el dato menor y el mayor en un conjunto de datos. Es una medida muy útil en cuestiones climatológicas pues muestra fielmente cuánto varió la temperatura en un lapso determinado de tiempo.

**Fórmula para calcular el rango:**

$$\text{RANGO} = \text{DATO MAYOR} - \text{DATO MENOR}$$

**Varianza** cuyo símbolo es ( $s^2$ ) es la media de las desviaciones al cuadrado, calculada usando **n** o **n-1** como divisor, dependiendo si es varianza poblacional o muestral respectivamente. Su expresión es la siguiente:

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

**Desviación típica o estándar** cuyo símbolo es (S) La desviación estándar es simplemente la raíz cuadrada positiva de la varianza. Su expresión es:



**Calcular la media aritmética, mediana, moda, rango, varianza y desviación estándar de los siguientes datos no agrupados 46,55, 50,47,52 50**

$$\text{Media aritmética} = \frac{46+55+50+47+52+50}{6} = \frac{300}{6} = 50$$

$$\text{Mediana} = 46, 47, 50, 50, 52, 55,$$

$$\text{Mediana} = 50$$

$$\text{Moda} = 50$$

$$\text{Rango} = 55 - 46 = 9$$

La media aritmética de 6 elementos se sabe que es 10. Sabiendo que cinco de ellos son: 8, 12, 13, 5 y 9, hallar el elemento que falta.

Calcular la media aritmética, mediana, moda, rango, varianza y desviación estándar de los siguientes datos no agrupados 46,55, 50,47,52 50

$$\text{Media aritmética} = \frac{46+55+50+47+52+50}{6} = \frac{300}{6} = 50$$

Para calcular la mediana lo que haremos primero es ordenar los datos de menor a mayor

$$\underline{46,47,50,50,52,55},$$

**Mediana=50**

**Moda=50**

Calcular varianza y desviación estándar de los siguientes datos no agrupados 46,55, 50,47,52 50

$$\begin{aligned}
 (s^2) &= \frac{(46-50)^2 + (55-50)^2 + (50-50)^2 + (47-50)^2 + (52-50)^2 + (50-50)^2}{6-1} = \\
 &= \frac{(-4)^2 + (5)^2 + (0)^2 + (-3)^2 + (2)^2}{5} = \\
 &= \frac{16+25+0+9+4+0}{5} = \frac{54}{5} = 10.8 \\
 &\quad \text{Varianza} = 10.8
 \end{aligned}$$

$$\sqrt{s^2} = \sqrt{10.8}$$

DESVIACION ESTANDAR=3.2

La desviación estándar es 6 cuanto es la varianza 36

La varianza es 36 la desviación estándar es 6

La **probabilidad** es una medida (comúnmente en la práctica expresada en %) que *muestra la proporción de veces con la que puede esperarse que ocurra cada uno de los* resultados de sucesos aleatorios con relación al total, donde cada resultado tiene la misma oportunidad de suceder (resultados equiprobables)

## PROPIEDADES DE LA PROBABILIDAD

### PROBABILIDAD DE UN EVENTO

Si se desea conocer la probabilidad de que suceda un evento se debe calcular la razón del número de posibles resultados que satisfacen la condición de este evento con respecto al número total de resultados igualmente posibles de ocurrir que componen el espacio muestral del fenómeno aleatorio.

$$P(A) = \frac{nA}{N}$$

Donde:

$nA$  = numero de resultado posibles del evento A.

$N$  = numero total de resultado en el espacio muestral S.

$P(A)$  = probabilidad de que suceda el evento A.

La probabilidad queda expresada en % después de multiplicar el cociente de por 100. Así, si la probabilidad de un evento es  $P(A) = 1$  entonces  $P(A) = 100\%$  y si  $P(B) = 0.5$  entonces  $P(B) = 50\%$ .

## Propiedades de probabilidad

La probabilidad de que suceda un evento A. Puede ser 0,1 o un número entre 0 y 1.

$$0 \leq P(A) \leq 1$$

La probabilidad de un espacio muestral S es 1.

$$P(S) = 1$$

La probabilidad de un evento que no puede ocurrir es 0

$$P(\emptyset) = 0$$

La probabilidad del complemento de un evento a (llamado A y que comprende Todas las respuesta que no se incluye en el resultado del evento) es  $1 - P(A)$

$$P(\bar{A}) = 1 - P(A)$$

# Experimento, espacio muestral evento

Aunque no podemos predecir el resultado de los fenómenos aleatorios, si es posible pronosticar “lo que es posible” realizando experimentos para provocar la repetición en condiciones similares de estos fenómenos.

## Experimento

Es un proceso o una acción que provoca Fenómenos aleatorios para observar y medir

## Espacio muestral

Es el conjunto de todos los resultados posibles de un experimento. Se Identifica con la letra  $S$  y su contenido se encierra entre  $\{ \}$ .

## Evento simple

Cualquier subconjunto del espacio muestral es un evento, puede ser uno de todos los resultados de un experimento o algunos de ellos que cumplan una condición

## Eventos compuesto

Se forma al combinar varios eventos simples. Si  $A$  y  $B$  son dos eventos, entonces:

- $A$  o  $B$
- $A$  y  $B$
- $A$  cuando sucede  $B$

Son eventos

compuesto

Obtener el espacio muestral en el lanzamiento de dos monedas

$$S = \{(a,a), (s,s), (a,s), (s,a)\}$$

$$2^2 = 4$$

OBTENER EL ESPACIO MUESTRAL EN EL LANZAMIENTO DE DOS DADOS AL MISMO TIEMPO

	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

$$6^2 = 36$$

Si un muchacho tiene en su guardarropa, 3 camisas color blanco, 2 azules, 4 camisas negras, 5 verdes y 2 camisas rojas y hoy para vestir elige una al azar ¿Cuál es la probabilidad de que se ponga una camisa color verde?

DATOS DEL PROBLEMA	FORMULA	SUSTITUCION	RESULTADO
<p><math>S=\{16\}</math></p> <p><math>A=\{\text{azules}\}=\{2\}</math></p> <p><math>N=16</math></p> <p><math>n(a)=2</math></p>	$P(A)=\frac{n(A)}{N}$	$P(A)=\frac{5}{16}=0.312$	<p>Existe una probabilidad de 31.25% de que elija la camisa azul.</p>

El muestreo probabilístico es un método de muestreo (muestreo se refiere al estudio o el análisis de grupos pequeños de una población) que utiliza formas de métodos de selección aleatoria.

**El requisito más importante del muestreo probabilístico es que todos en una población tengan la misma oportunidad de ser seleccionados.**

**Muestreo aleatorio simple:** es aquel en el que todos los individuos de la población tienen la misma probabilidad de ser elegidos.

**Cuando se usan estas tablas para seleccionar una muestra aleatorio simple debemos seguir los siguientes pasos:**

1. Enumerar todos los elementos de la población (N).
2. Elegir el tamaño de la muestra (n)
3. Elegir aleatoriamente las columnas a utilizar en la tabla de números aleatorios.
4. Tomar de manera consecutiva , ya sea de arriba hacia abajo o viceversa los números correspondiente elegibles , que caigan entre 1 y N
5. Una vez tomado los n números requeridos , estos comprenderán ala muestra aleatoria

Muestreo estratificado: este es un método en el cual una población grande se divide en dos grupos más pequeños, que generalmente no se superponen, sino que representan a toda la población en conjunto.

**El muestreo por conglomerados** por lo general analiza a una población particular en la que la muestra consiste en varios elementos, por ejemplo, ciudad, familia, universidad, etc. Los conglomerados se seleccionan básicamente dividiendo la población mayor en varias secciones más pequeñas.

**Muestreo sistemático:** este se enfoca en elegir a cada “enésima” persona para que sea parte de la muestra. Por ejemplo, puedes elegir que cada quinta persona sea parte de la muestra, o que cada décima persona sea parte de ella.

El muestreo **sistemático** es una implementación extendida de la mismísima técnica de probabilidad en la que cual, cada miembro de un grupo es seleccionado en periodos regulares para formar una muestra. Cuando se utiliza este método de muestreo, existe una oportunidad igual para que cada miembro de una población sea seleccionado.

Con las siguientes 20 edades tomado de un grupo de estudiantes de la UDS obtener mediante el **muestreo aleatorio simple** una muestra de tamaño  $n=5$  las edades son las siguientes: 18, 19, 20, 21, 22, 23, 24, 25, 18, 19, 20, 21, 22, 23, 24, 25, 26, 25, 24, 23

**$N=20$ , N, población**

**$n=5$ , n, muestra**

**Ran# = Numero aleatorio**

**$20 \text{Ran\#} = 19.56 = 20$**

**$20 \text{Ran\#} = 11.6 = 12$**

**$20 \text{Ran\#} = 7.78 = 8$**

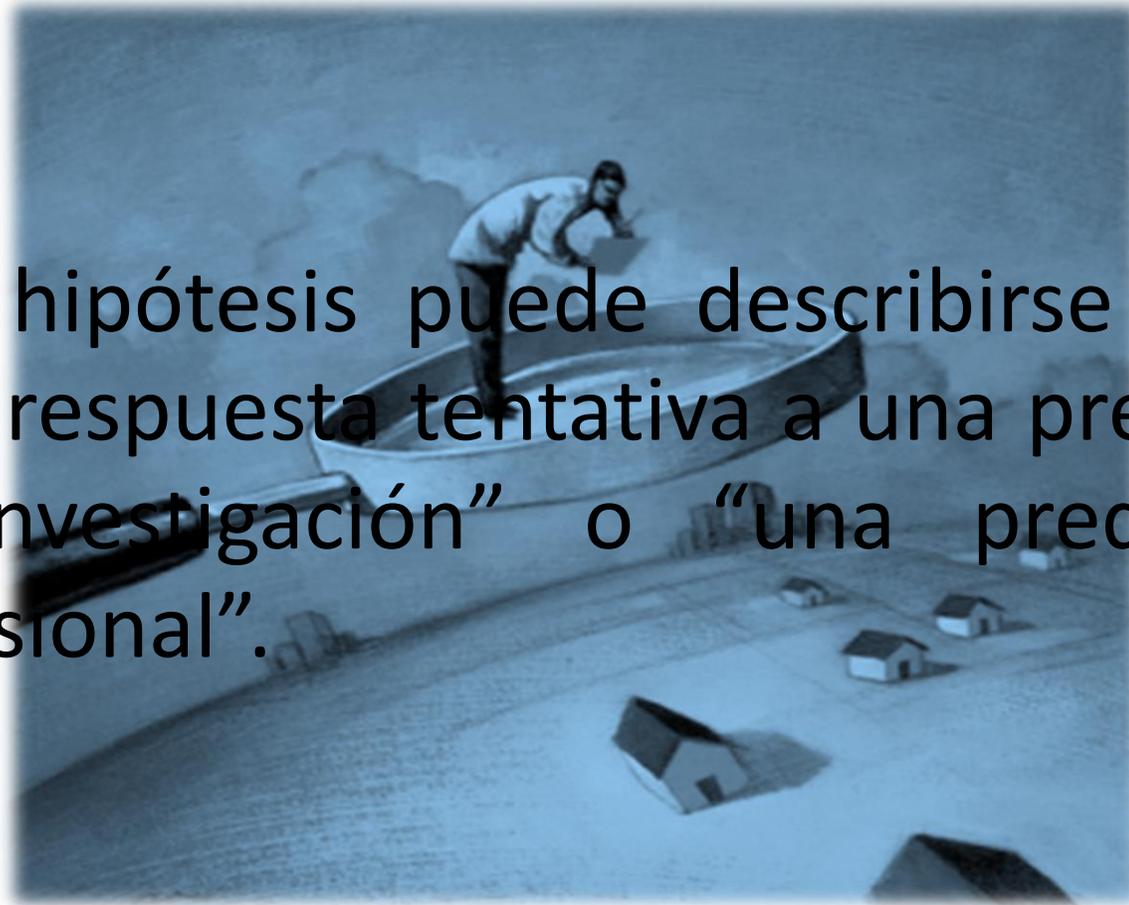
**$20 \text{Ran\#} = 18.52 = 19$**

**$20 \text{Ran\#} = 3.3 = 3$**

**La muestra seleccionada es: 23, 21, 25, 24, 20**

# 1. Hipótesis y Variables

Una hipótesis puede describirse como “una respuesta tentativa a una pregunta de investigación” o “una predicción provisional”.



## La hipótesis debe ser:

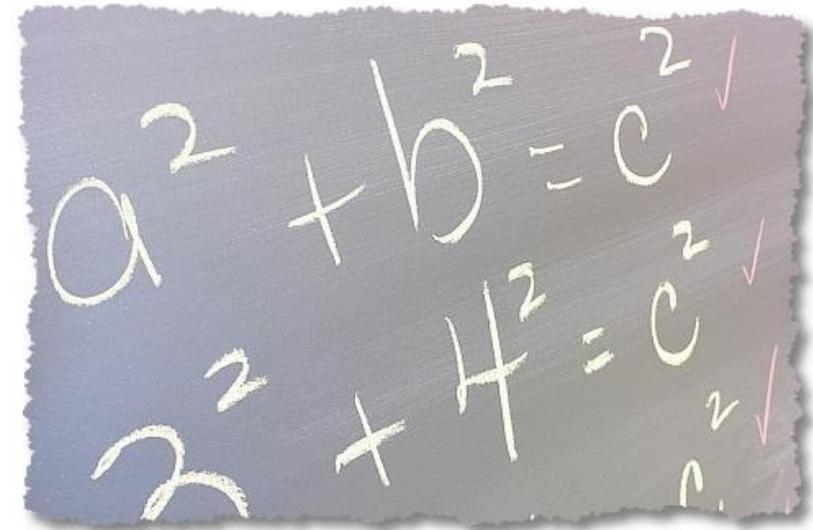
- Claramente enunciada, utilizando terminología apropiada
- Comprobable
- Una declaración de relaciones entre variables
- Limitada en su alcance



- Los programas de educación para la salud influyen sobre el número de gente que fuma
- Los periódicos afectan los patrones de votación de la gente
- La asistencia a clases influye en las calificaciones
- La dieta influye sobre la inteligencia

En estos ejemplos “algo” (por ejemplo, la dieta la asistencia a clase) afecta a “algo mas” (por ejemplo, la inteligencia, las calificaciones). Estas son **variables**.

Una variable es algo que se encuentra libre para variar, y para describirlas **cuantitativamente**, tienen que ser expresadas en unidades apropiadas (por ejemplo, Nivel de IQ, calificaciones numéricas).



Los pares de variables en los ejemplos dados tienen nombres distintos.

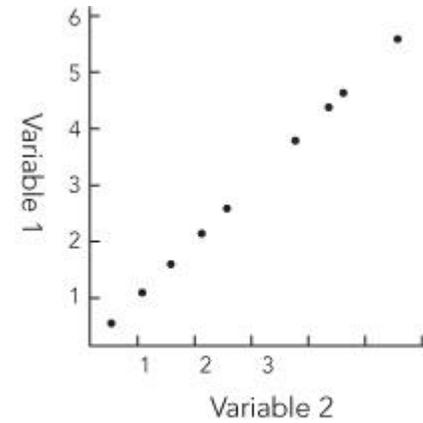
La variable que manipulamos se llama **variable independiente (VI)**.

La variable con la que estamos haciendo la hipótesis que cambiará como resultado de la manipulación se llama **variable dependiente (VD)**.

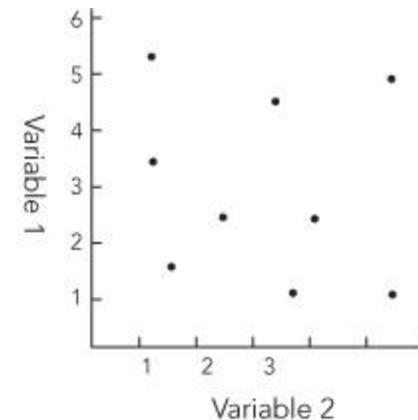


✱ La variable dependiente cambia como consecuencia del valor de la variable independiente – su valor depende de esto.

✱ El valor de la variable independiente es libre para variar de acuerdo a los caprichos del experimento.



Perfect positive correlation  
Correlation coefficient = +1



No correlation  
Correlation coefficient = 0

VARIABLE INDEPENDIENTE	VARIABLE DEPENDIENTE
Programas de educación sanitaria	Número de gente que fuma
Periódico	Patrones de votación
Asistencia a clase	Calificaciones en los exámenes
Dieta	Inteligencia

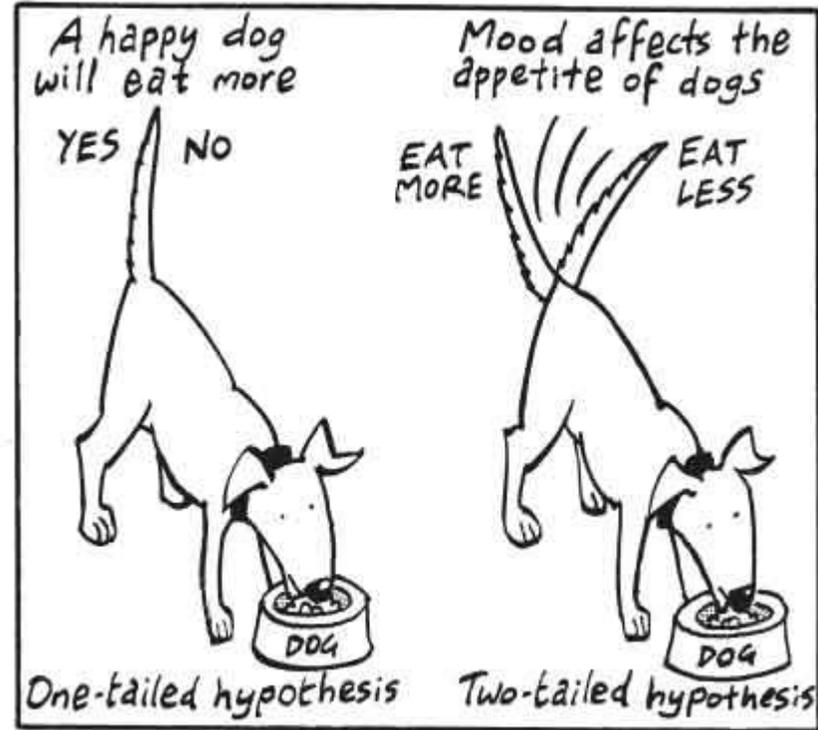
Muchas variables pueden ser tanto dependientes como independientes dentro del contexto de un estudio en particular. Por ejemplo, se puede sostener que *“la inteligencia influye sobre la dieta”* o que *“las calificaciones en los exámenes influyen sobre la asistencia a clase”*.

## a) Hipótesis direccional

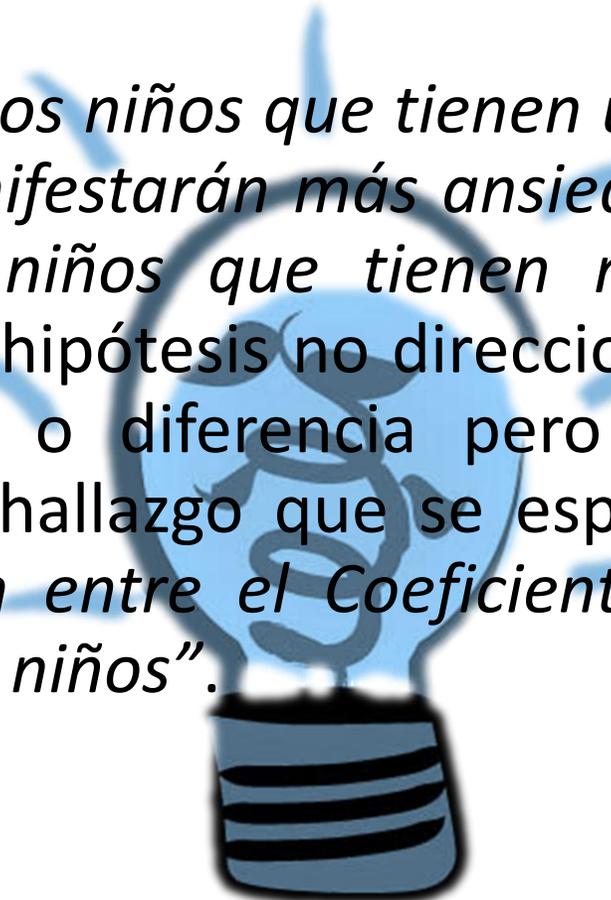
En los ejemplos anteriores palabras como “influye” o “afecta” son utilizadas sin indicar dirección.

Cuando una hipótesis afirma un resultado predicho utilizando palabras como reduce, incrementa, baja, alza – se le llama hipótesis direccional o hipótesis de una cola.

Los tipos de hipótesis mas vagas son conocidas como no direccionales o de dos colas. La hipótesis direccional especifica la naturaleza de la relación o la diferencia pronosticada.



Por ejemplo: *“Los niños que tienen un alto Coeficiente Intelectual manifestarán más ansiedad en el salón de clase que los niños que tienen menor Coeficiente Intelectual”*. La hipótesis no direccional establece que existe relación o diferencia pero no especifica la naturaleza del hallazgo que se espera. Por ejemplo: *“Existe relación entre el Coeficiente Intelectual y la ansiedad en los niños”*.

A stylized illustration of a glowing lightbulb with a blue glow and radiating lines, symbolizing an idea or insight. The lightbulb is positioned behind the text, with its base at the bottom center.



**El concepto de salud que se adopta es el que la Organización Mundial de la Salud (OMS) elaboró en 1947, que afirma que la salud es “un estado de completo bienestar físico, mental y social, y no solamente la ausencia de afecciones o enfermedades”.**

El verbo “medir” se refiere al procedimiento de aplicar una escala patrón a una variable o un conjunto de variables, mientras que el sustantivo “medición” se refiere a la extensión, dimensión, cantidad, etc., de un atributo. Según Morgenstern, medir variables de salud abarca diferentes niveles de medición, que pueden generarse de dos maneras:

1. Por la observación directa de la persona (por ejemplo, la presión arterial de las personas, el acceso de las personas a los servicios de salud cuando los necesitan).

2) Por la observación de un grupo poblacional o un lugar mediante tasas y proporciones (como la tasa de prevalencia de hipertensión o el porcentaje de adolescentes entre 15 y 19 años de edad que son madres), medias (media de consumo de sal per cápita en un municipio) y medianas (mediana de supervivencia de enfermos con cáncer), entre otras. Las mediciones generadas a partir de la observación de grupos o lugares se usan para generar indicadores y pueden clasificarse de la siguiente manera:

Algunos ejemplos de tasas son: la tasa de natalidad, que es la relación de los nacidos vivos al número de habitantes durante un año; tasas de interés que expresan la cantidad de dinero que una inversión produce durante un plazo determinado, etcétera. Si una de las variables es el tiempo, la tasa se denomina tasa de cambio; por ejemplo, la velocidad de un automóvil, que es la tasa de la distancia recorrida al tiempo invertido en el recorrido, o el cambio en el nivel de agua al llenar una alberca. En problemas específicos se usan la tasa de fecundidad, tasa de mortalidad, tasa de inmigración, tasa de divorcio, tasa de crecimiento, etcétera.

Tasa: es un indicador que mide el riesgo de enfermar, morir o sufrir daño a la salud en determinado tiempo.

$$\text{Tasa} = \frac{\text{Numero de sucesos, eventos favorables}}{\text{Total de la poblacion}} 10^n$$



$$\text{Tasa} = \frac{\text{Numero de sucesos, eventos favorables}}{\text{Total de la poblacion}} 10^n$$

Ejemplo: En México en el año 2013 se registraron 2 162 535 nacimientos de una población de 119 530 753 determinar la tasa de natalidad.

$$\text{Tasa de natalidad} = \frac{2162535}{119530753} (10000) = 180$$

Concluimos que la tasa de natalidad fue de 180 nacidos vivos por cada 10,000 habitantes.

## RAZONES

La mayor parte de la información que procesamos todos los días se basa en la relación de cantidades que expresamos como fracciones, razones, proporciones o porcentajes. Un alumno sabe que una medida como el promedio de sus calificaciones informa sobre su estado de aprendizaje o que un porcentaje expresa la cantidad de una población que tiene ciertas características; por ejemplo, el porcentaje de alumnos que juega ajedrez en tu escuela.

**Una *razón* es la relación de dos cantidades para expresar cuánto de una está contenida en (o pertenece a) la otra. La notación empleada para expresar esta relación es  $a:b$ , que se lee  $a$  es a  $b$ .**

Por ejemplo, si en un salón hay 36 mujeres y 24 hombres, la razón de mujeres a hombres es de 36 a 24. En nuestro ejemplo, la razón de mujeres a hombres en el salón es **36: 24** . La expresión  **$a+b$**  es la cantidad total y  **$a$  y  $b$**  son las partes del total que se relacionan. En realidad, tratamos de saber cuántas mujeres hay por cada hombre en el salón, de modo que está implícita la operación de división en esta relación; así, **36: 24** es lo mismo que:

$$\frac{36}{24} = \frac{18}{12} = \frac{9}{6} = \frac{3}{2}$$

**Podemos concluir que hay 3 mujeres por cada 2 hombres o que el 60% son mujeres por lo tanto el 40% son hombres.**

## PROPORCIÓN:

En ocasiones disponemos de dos razones  $a:b$  y  $c:d$  por ejemplo, las razones de mujeres a hombres en dos salones diferentes; las razones de altura a longitud en dos pizarrones; las razones de hembras a machos en dos peceras; etcétera

Una proporción es la igualdad entre dos razones. La expresión de una proporción es  $a:b :: c:d$  Que también se puede escribir como:  $\frac{a}{b} = \frac{c}{d}$

**Ejemplo 1:** En un salón hay 36 mujeres y 24 hombres, ¿cuántas mujeres debe haber en otro salón que tiene 18 hombres para que los grupos sean proporcionales?

La razón en el primer grupo, calculada anteriormente, es **3:2** , y la razón en el segundo salón es  $c:18$ . Para que los grupos sean proporcionales, se debe cumplir que **3:2 :: c:18** ; es decir:

$$\frac{3}{2} = \frac{c}{18}$$