	INSTITUCIÓN EDUCATIVA "EL RECUERDO" Resolución de Aprobación de Carácter Oficial No. 0143 de 2017 en los niveles de Preescolar, Básica y Media Académica DANE. 123001800064 NIT. 901048820-9	Fecha

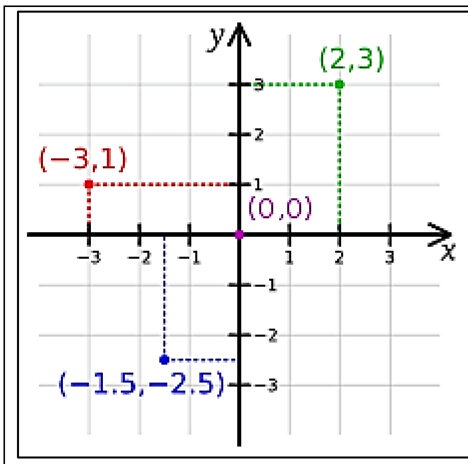
Guía de trabajo del área: Estadística – Guía 5	**Grado:** 9A - 9B

Nombre del docente 9A: Ureliano Peñata **email:** upenataieelrecuerdo@gmail.com **Celular:** 3135276620

Nombre del docente 9B: Rosa Cano **email:** rcanoieelrecuerdo@gmail.com **Celular:** 3105679770

TEMAS Y/O SABER	DBA (APRENDIZAJES)
✓ **Variables estadísticas bidimensionales** ✓ **Tablas de doble entrada** ✓ **Dependencia aleatoria y funcional** ✓ **Diagramas de dispersión**	**DBA 10:** Propone un diseño estadístico adecuado para resolver una pregunta que indaga por la comparación sobre las distribuciones de dos grupos de datos, para lo cual usa comprensivamente diagramas de caja, medidas de tendencia central, de variación y de localización.

**********Nota: La mejor manera de trabajar tu guía es la siguiente:*** *Inicia por la sección RECORDEMOS O SABERES PREVIOS y realiza las actividades allí propuestas. Luego continúa con la sección APRENDAMOS, transcribe todo el contenido de esta sección en tu cuaderno, leyendo los ejemplos, comprendiéndolos y resolviéndolos en tu cuaderno. En caso que no comprendas un concepto o ejemplo escribe a tu docente a los números que están en el encabezado de esta guía según el horario de clases establecido. Luego de comprender los temas, ya puedes proceder a resolver los talleres. Envía al correo electrónico los talleres resueltos de la sección PRACTIQUEMOS, preferiblemente en formato PDF indicando Asignatura, número de guía y nombre del estudiante. Solo envía el taller* ***********



SABERES PREVIOS

Recuerda cómo se ubica en el plano cartesiano una pareja ordenada. Recuerda que en el plano cartesiano la primera coordenada corresponde a la variable *x* y la segunda, a la varibale *y.* Observa los ejemplos de los puntos ubicados en el plano cartesiano que se muestra a continuación. Analiza y comprende la ubicación de los mismos, recorta y pega o dibuja este diagrama en tu cuaderno.

Un error comun es ubicar cada punto en el eje equivocado. Por ejemplo en el punto (2,3) se debe ubicar el dos en el eje de las X y en 3 en el eje de las Y.

APRENDAMOS

Variables estadísticas bidimensionales

Las variables que se obtienen al observar simultáneamente dos características de un mismo elemento de una población estadística se llaman **variables estadísticas bidimensionales.** Una **variable bidimensional** es una **variable** en la que cada individuo está definido por un par de caracteres, (X, Y). Estos dos caracteres son a su vez **variables estadísticas** en las que sí existe relación entre ellas, una de las dos **variables** es la **variable** independiente y la otra **variable** dependiente. Estas se representan mediante el par (X, Y) y toman los valores (X_1, Y_1); (X_2, Y_2); … (X_n, Y_n).

<u>Ejemplo</u>: La siguiente tabla muestra algunos ejemplos de variables bidimensionales:

(X,Y)	X	Y
(sexo, color del pelo)	*cualitativo*	*cualitativo*
(profesión, antigüedad en la empresa)	*cualitativo*	*cuantitativo*
(peso, estatura)	*cuantitativo (v.e. continua)*	*cuantitativo (v.e. continua)*
(número de hermanos, número de hijos)	*cuantitativo (v.e. discreta)*	*cuantitativo (v.e. discreta)*
(temperatura, pulsaciones)	*cuantitativo (v.e. continua)*	*cuantitativo (v.e. discreta)*

Se habla de variables bidimensionales cuando interesa estudiar simultáneamente dos (o más) caracteres de una población. Si de una cierta población se estudian dos caracteres simultáneamente se obtienen dos series de datos.

<u>Ejemplo:</u> Entre los empleados de una empresa se ha realizado una encuesta sobre el consumo del tabaco, que ha arrojado los siguientes resultados:

Hábito / Sexo	Fumadores	No fumadores	Totales de filas
Varones	49	64	113
Mujeres	43	37	80
Totales de columnas	92	101	Total general **193**

Nota. En este tema nos limitaremos al estudio de caracteres cuantitativos discretos, puesto que si el carácter es continúo o discreto agrupado en intervalos, se trabajará con las marcas de clase.
A continuación, estudiaremos el ejemplo de la sección ANALIZA y CONOCE que está en la página 116 del texto guía.

ANALIZA

Se preguntó a 50 personas acerca del uso del transporte público a lo largo de un mismo día y se registraron los datos en la Tabla. ¿Cuántas personas toman taxi 2 veces?, ¿cuántas viajan en autobús en dos ocasiones?, ¿cuántas usan taxi y autobús dos veces al día?.

Taxi (X) / Autobús (Y)	0	1	2	Total
0	2	2	10	14
1	4	8	8	20
2	8	6	2	16
Total	14	16	20	50

CONOCE

Si se tiene en cuenta que en la primera fila están ubicados los valores de la variable X: “número de veces que se usa taxi” y en la primera columna los de la variable Y: “número de veces que se usa autobús”, puede decirse que:

- Hay 20 personas que toman taxi dos veces.
- Hay 16 personas que viajan en autobús en dos ocasiones.
- Hay dos personas que usan taxi y autobús dos veces al día

Tablas de doble entrada

La tabla de doble entrada se utiliza cuando el número de valores observados de una variable bidimensional es bastante grande y se repiten muchos de ellos.

Ejemplo: A continuación se muestra una tabla de doble entrada en la que se figuran los resultados de una encuesta hecha a los empleados de una empresa para estudiar la incidencia del tabaquismo en la gravedad de los accidentes laborales.

	Muy grave	Grave	Lesiones medianas	Leves
Muy fumador	20	10	10	30
Fumador	30	40	20	50
Fumador esporádico	10	60	80	60
No fumador	5	20	30	50

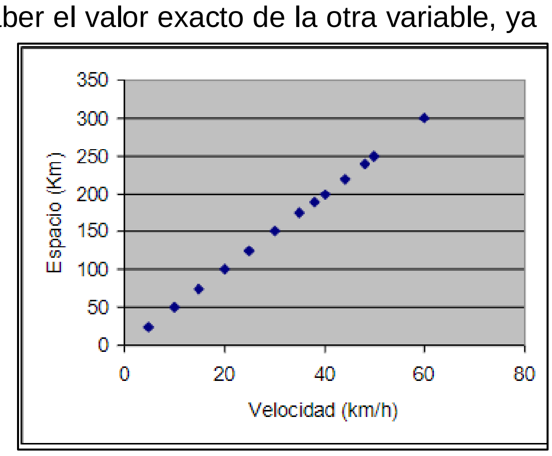
Con base en los resultados, se puede afirmar que de los trabajadores muy fumadores sólo 10 han sufrido accidentes laborales graves; en cambio, de los fumadores esporádicos, 60s al lesionado gravemente.

Nota: Revisa el ejemplo para que evidencies lo que se concluye en él.

Dependencia aleatoria y funcional

Entre dos variables podemos establecer una relación, la cual puede ser:

- ✓ **Dependencia funcional:** Recordemos que la dependencia funcional la hemos estudiado en las clases de matemáticas, por ejemplo cuando una variable X y una variabe Y, se r elacionan entre sí y dependen una de la otra. En algunos estudios donde se analizan dos variables observamos que, conociendo el valor de una de las variables podemos saber el valor exacto de la otra variable, ya que hay una formula matemática que las relaciona. En la vida diaria podemos buscar ejemplos de estas dependencias funcionales, como es el caso de la relación entre espacio recorrido y velocidad de un automóvil, para un intervalo de tiempo fijo. Como vemos en la tabla, dando valores distintos al espacio para un intervalo fijo de 5 horas, nos da como resultado un único valor para cada valor del espacio, de forma exacta, pues las variables están relacionadas por la ecuación:



$$V = \frac{e}{t};\ V = \text{velocidad};\ e = \text{espacio};\ t = \text{tiempo}$$

Velocidad (Km/h)	5	10	15	20	25	30	35	38	40	44	48	50	60
Espacio (Km)	25	50	75	100	125	150	175	190	200	220	240	250	300

En el gráfico se puede apreciar el comportamiento de la relación existente entre las variables Espacio y Velocidad. Esta relación es directa, en la medida que incrementa una tambien incrementa la otra y su relación está expresada con la ecuación anteriormente expuesta.

- ✓ **Dependencia aleatoria o correlación:** En este caso los valores no se ajustan a la gráfica de una función, pero guardan cierta relación. Existen casos en los que no se puede establecer con exactitud una formula que relacione dos variables, esto no quiere decir que no exista una cierta relación. Por ejemplo si queremos relacionar el coeficiente de inteligencia con las notas obtenidas en un examen, es lógico pensar que cuanto más alto es el coeficiente de inteligencia la nota será mejor, aunque esto no ocurra en todos los casos o pueda variar ligeramente. Este tipo de dependencia llamaremos dependencia aleatoria.

Si los valores de ambas variables los representamos en el plano cartesiano, los puntos, en general, no se ajustan de un modo preciso a una función matemática, sino que se obtiene un conjunto de puntos más o menos dispersos. Una representación de ese tipo recibe el nombre de <u>nube de puntos</u> o <u>diagrama de dispersión</u>.

El **<u>diagrama de dispersión</u>** o la nube de puntos se utiliza para tener una idea de la relación que existe entre dos variables. La relación será más intensa cuanto más concentrados estén los puntos y puede ser directa (si cuando al aumentar el valor de una variable aumenta el valor de la otra) o inversa (si cuando al aumentar una variable disminuye el valor de la otra).

Observa las figuras 4.22 a 4.24.

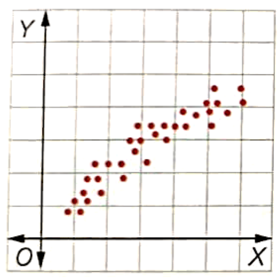


Figura 4.22

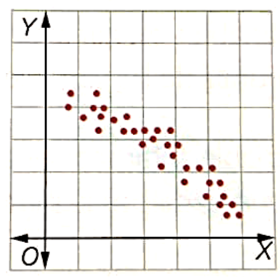


Figura 4.23

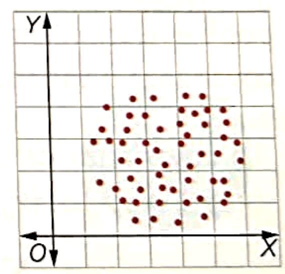


Figura 4.24

- La nube de puntos de la Figura 4.22 es muy estrecha; por tanto, hay una dependencia significativa y al aumentar una variable, aumenta la otra, es decir, presentan una **dependencia** o una **correlación positiva o directa**.
- La nube de puntos de la Figura 4.23 es muy estrecha; por consiguiente, también hay una dependencia significativa y al aumentar una variable, la otra disminuye; esto es, presentan una **dependencia** o una **correlación negativa o inversa**.
- En la Figura 4.24 no se observa ninguna dependencia entre las variables; por ello, se dice que su **correlación es nula**.

Ejemplo

En la Tabla 4.32, se muestran la estatura, en centímetros, de diez personas y el número de calzado que usan.

Estatura (cm)	154	156	157	158	160	163	165	170	171	172
Número de calzado	35	36	36	36	37	38	38	40	41	41

Tabla 4.32

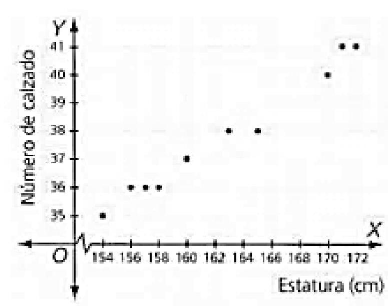


Figura 4.25

Al representar los pares de valores de la variable estadística bidimensional (Estatura, Número de calzado), se obtiene un conjunto de puntos que representan una correlación positiva, como se observa en la Figura 4.25.

<u>Cómo hacer un diagrama de dispersión paso a paso</u>

- **Paso 1**: Determina cuál es la situación. Si no entendemos qué es lo que esta ocurriendo, no podremos establecer las variables a estudiar.
- **Paso 2**: Determina las variables a estudiar. Si ya determinaste las variables a estudiar, es porque crees que puede existir una relación entre ellas que te permita caracterizar la situación.
- **Paso 3**: Recolecta los datos de las variables: Si ya los tienes, perfecto. Si no, definimos un período de tiempo para conseguir los datos de las variables antes definidas. Recuerda que los datos de las dos variables deben estar dados en el mismo período de tiempo.
- **Paso 4**: Ubica los valores en el eje respectivo. Por lo general, la variable independiente es aquella que no está influenciada por la otra y se ubica en el *eje x*. La variable dependiente que es la que se ve afectada por la otra variable se ubica en el eje y. Así pues, procedemos a ubicar los valores en el plano cartesiano de acuerdo a su variable (x, y)

- **Paso 5**: Determina el coeficiente de correlación: El coeficiente de correlación debe verse reflejado en la forma que toma el gráfico de dispersión. Es el cociente de la covarianza y la multiplicación de la desviación típica de las dos variables. Con excel logramos calcularlo de manera muy simple.
- **Paso 6**: Analizamos: Con base en el coeficiente y en el gráfico, definimos cuál es la relación de las dos variables y tomamos las decisiones pertinentes.

Algunos ejemplos de posibles tipos de relación que pueden mostrar los diagramas de dispersión son los siguientes:

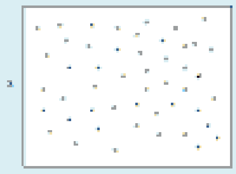
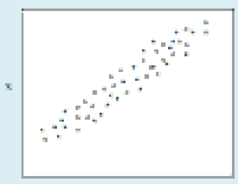
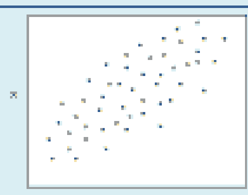
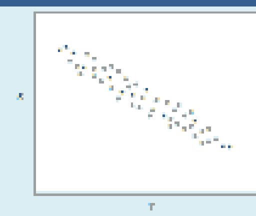
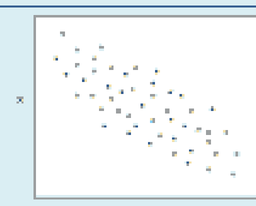
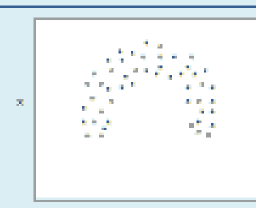
Diagrama	Tipo de Relación
	Sin relación. No se aprecia ninguna correlación entre las dos variables.
	Alta correlación positiva. El valor de Y se incrementa nitidamente a medida que el valor de X aumenta.
	Baja correlación positiva. El valor de X aumenta ligeramente a medida que aumenta el valor de Y.

Diagrama	Tipo de Relación
	Fuerte correlación negativa. El valor de X claramente disminuye a medida que aumenta el valor de Y.
	Débil correlación negativa. El valor de X disminuye ligeramente a medida que aumenta el valor de Y.
	Relación compleja. El valor de X parece estar relacionado con el valor de Y, pero esa relación no es fácil de establecer.

Correlación lineal

La correlación o coeficiente de correlación lineal es una medida estadística que cuantifica la dependencia lineal entre dos variables, es decir, si se representan en un diagrama de dispersión los valores que toman dos variables, el coeficiente de correlación lineal señalará lo bien o lo mal que el conjunto de puntos representados se aproxima a una recta.

Si la representación de la relación entre dos variables es una nube de puntos que se condensa en torno a una recta, entonces existe una **correlación lineal** entre las variables. El **coeficiente de correlación lineal** se define como:

$$r = \frac{S_{xy}}{S_x S_y} \text{ y está comprendido entre } -1 \text{ y } 1.$$

Valores que puede tomar la correlación

r = -1 Correlación perfecta negativa

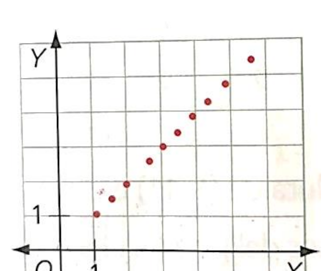
r = 0 No existe correlación

r = +1 Correlación perfecta positiva

Hablamos de correlación positiva si siempre que el valor «x» sube, el valor «y» sube, y además con la misma intensidad (+1).

En el caso opuesto, si siempre que el valor «x» sube, y el valor «y» baja, y además con la misma intensidad, entonces estamos hablando de correlación negativa (-1).
Es importante saber que esto no quiere decir que lo hagan en la misma proporción (salvo que tengan la misma desviación típica).

Responde y comparte la pregunta de la sección ANALIZA de tu texto guía:
En la figura se representan dos variables correlacionadas. ¿Qué tipo de correlación hay entre las variables representadas?



Ejemplo:

Observa la relación que hay entre el diagrama de dispersión y el coeficiente de correlación lineal en los casos que se muestran en las figuras 4.27 a 4.31.

$r = -1$

Correlación negativa y perfecta. Los puntos están alineados. Dependencia funcional.

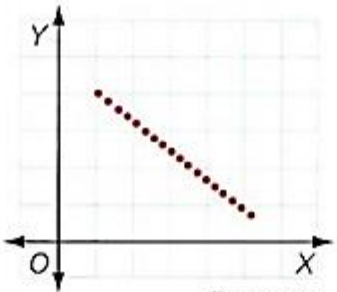


Figura 4.27

$-1 < r < 0$

Correlación negativa, más fuerte cuanto más se aproxima r a -1, y más débil cuanto más se aproxima a 0. Dependencia aleatoria.

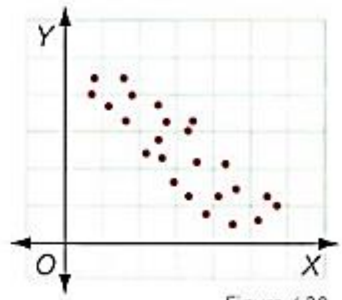


Figura 4.28

$0 < r < 1$

Correlación positiva, más fuerte cuanto más se aproxima r a 1 y más débil cuanto más se aproxima a 0. Dependencia aleatoria.

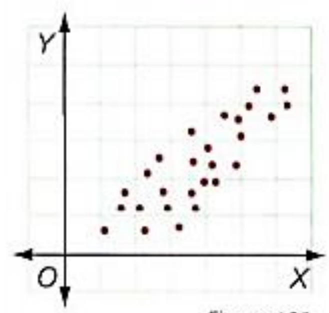


Figura 4.29

$r = 1$

Correlación positiva y perfecta. Los puntos de la nube están alineados. Dependencia funcional.

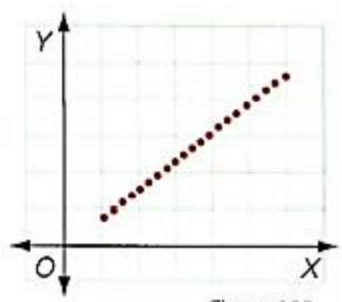
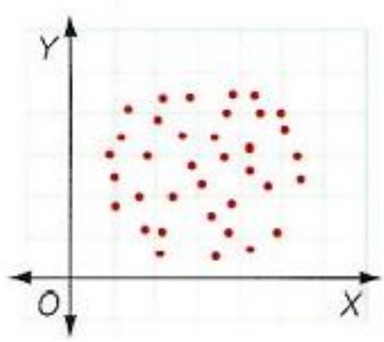


Figura 4.30

Si r = 0 las variables no tienen correlación, es decir, hay una independencia aleatoria. Un ejemplo de ellos lo encontramos en la figura que se muestra continuación.



"ASESORIA: si tiene alguna duda o no entiende algo sobre esta guía, comuníquese con el número que aparece en la parte de arriba"

PRACTIQUEMOS

Resuelve los ejercicios y envíalos al correo electrónico que aparece en el encabezado de la guía:

1. Completa la tabla:

Y \ X	A	B	C	Total
a	2		1	4
b		2	2	5
c			0	7
Total	7	6		

3. Observa la variable bidimensional registrada en la tabla y realiza lo que se pide.

Número de cigarrillos consumidos al día	3	6	8	20	25
Índice de mortalidad %	0,2	0,4	0,5	1,2	1,7

Tabla 4.36

a. Representa la nube de puntos.

b. Indica el tipo de correlación.

2. Indica qué tipo de relación tienen las variables bidimensionales (X, Y_1), (X,Y_2) y (X, Y_3). Nota: Deberás indicar si la correlación existente es positiva, negativa o nula.ten en cuenta además que la correlación puede ser débil o fuerte, indica esto tambien al responder el ejercicio.

X	23	21	2	4	5	7
y_1	9	5	21	25	27	211
y_2	4	3	2	1	0	21
y_3	22	2	21	1	5	3

4.

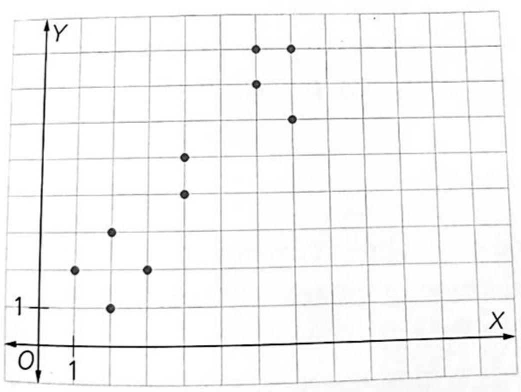
Durante su primer año de vida, han pesado a Miranda cada mes. En la Tabla 4.40, aparecen sus pesos. X representa la edad (meses) y Y, el peso (kilogramos).

X	1	2	3	4	5	6	7	8	9	10	11	12
Y	3,2	3,7	4,2	5,3	5,7	6,5	6,8	7,2	7,9	7,7	8	8,5

Tabla 4.40

a. Representa el diagrama de dispersión.

5. Obsetva el diagrama de dispersión y luego resuelve:



a. Elabora una tabla de doble entrada para registrar los datos de las variables X y Y.

b. ¿Qué tipo de correlación tienen las dos variables? ¿Fuerte o débil?, ¿positiva o negativa?

PROFUNDIZA: Si no cuentas con el texto guía lo puedes descargar en el siguienten link: https://tecevolucion.files.wordpress.com/2018/01/matematicas-9-vamos-a-aprender.pdf