Tabla de contenido

1. IDENTIFICACIÓN: 2
Competencias 2
Resultados de aprendizaje 2
2. PRESENTACIÓN: 2
3. UNIDADES DE APRENDIZAJE: 2
UNIDAD 1: VARIABLE ESTADISTICA 2
TIPOS DE VARIABLE ESTADÍSTICAS 2
VARIABLE CUANTITATIVA 2
ACTIVIDAD 1 3
UNIDAD 2: POBLACION ESTADISTICA 3
ACTIVIDAD 2 3
MEDIDAS DE POSICIÓN 3
ACTIVIDAD 3 4
ACTIVIDAD 4 5
ACTIVIDAD 5 6
UNIDAD 3: MEDIDAS DE DISPERSIÓN 6
RANGO, AMPLITUD TOTAL O RECORRIDO 7
DESVIACIÓN MEDIA 7
ACTIVIDAD 6 8
VARIANZA 8
DESVIACIÓN ESTANDAR 10
ACTIVIDAD 7 11
UNIDAD 4: REPRESENTACIONES GRAFICAS 12
ACTIVIDAD 8 14
ACTIVIDAD 9 19
UNIDAD 5: DISTRIBUCION DE PROBABILIDAD DISCRETA 19
ACTIVIDAD 10 25
ACTIVIDAD 11 25
ACTIVIDAD 12 25
ACTIVIDAD 13 25
ACTIVIDAD 13 25
4. GLOSARIO: 25
5. REFERENTES BILBIOGRAFÍCOS: 26
6. CONTROL DEL DOCUMENTO: 26
7. CONTROL DE CAMBIOS: (diligenciar únicamente si realiza ajustes a la guía) 26

1. IDENTIFICACIÓN:

Área: Matemáticas (Estadística) **Grado:** Décimo **Tiempo:** 10 meses

Competencias

Identifica los conceptos básicos de la estadística descriptiva.
Categoriza las diferentes variables según el problema planteado.
Calcula las medidas de centralización, localización y dispersión de los datos dados.

Resultados de aprendizaje

Comprensión de las medidas de centralización, localización y dispersión, interpretando cada resultado numérico para traducirlo al lenguaje cotidiano.
Construcción y organización de las muestras.
Reconocimiento de la tendencia de las muestras a lo largo de un intervalo de tiempo.

2. PRESENTACIÓN:

Esta guía está diseñada para el desarrollo de las habilidades enfocado al estudio de las medidas de centralización, localización y dispersión. Presentando con gran detalle el significado de la media, desviación media y la desviación estándar.

3. UNIDADES DE APRENDIZAJE:

UNIDAD 1: VARIABLE ESTADISTICA

Una variable estadística es cada una de las características o cualidades que poseen los individuos de una población.

TIPOS DE VARIABLE ESTADÍSTICAS

- **Variable cualitativa**

Las variables cualitativas se refieren a características o cualidades que no pueden ser medidas con números. Podemos distinguir dos tipos:

- **Variable cualitativa nominal**

Una variable cualitativa nominal presenta modalidades no numéricas que no admiten un criterio de orden. Por ejemplo:
El estado civil, con las siguientes modalidades: soltero, casado, separado, divorciado y viudo.

- **Variable cualitativa ordinal**

Una variable cualitativa ordinal presenta modalidades no numéricas, en las que existe un orden. Por ejemplo:
La nota en un examen: Deficiente, Insuficiente, Aceptable, sobresaliente, excelente.
Puesto conseguido en una prueba deportiva: 1º, 2º, 3º…
Medallas de una prueba deportiva: oro, plata, bronce.

VARIABLE CUANTITATIVA

Una variable cuantitativa es la que se expresa mediante un número, por tanto se pueden realizar operaciones aritméticas con ella. Podemos distinguir dos tipos:

- **Variable discreta**

Una variable discreta es aquella que toma valores aislados, es decir no admite valores intermedios entre dos valores específicos. Por ejemplo:
El número de hermanos de 5 amigos: 2, 1, 0, 1, 3.

- **Variable continúa**

Una variable continua es aquella que puede tomar valores comprendidos entre dos números. Por ejemplo:
La altura de los 5 amigos: 1.73, 1.82, 1.77, 1.69, 1.75.

ACTIVIDAD 1

1. Clasifica las siguientes variables en cuantitativa discreta o continua
   a. Número de Hijos de una familia
   b. Temperatura registrada cada hora en un observatorio
   c. Numero de hojas de vida que llega a una bolsa de empleo por día.
   d. El peso de 50 niños de un jardín

2. Clasifica las siguientes variables estadísticas en discretas y continuas.
   a. Número de acciones vendidas cada día en la Bolsa.
   b. Período de duración de un automóvil.
   c. El diámetro de las ruedas de varios coches.
   d. Número de litros de agua contenidos en un depósito.
   e. Suma de puntos tenidos en el lanzamiento de un par de dados.

UNIDAD 2: POBLACION ESTADISTICA

Es el conjunto de datos observables que poseen una característica determinada.

MUESTRA: Es un subconjunto de la población que debe ser representativa.

Tamaño poblacional (N): Es el número total de elementos de la población.

Tamaño muestral (n): Número de elementos de la muestra.

ACTIVIDAD 2

1. Identificar la población y la muestra en el siguiente estudio

a. El departamento de educación física de un colegio decide implementar actividades deportivas extra clases a los estudiantes que presentan bajo rendimiento en el área. Para ello, van a trabajar con los estudiantes que sean autorizados por los padres.
b. Se desea realizar un estudio en los hogares de Medellín para determinar cuál es su nivel de información y prevención frente al virus gripal AH1N1, esto se realizará por comunas.
c. El ministerio de protección social desea determinar los índices de mortalidad infantil en Medellín. Para ello va a realizar un estudio del número de casos registrados en algunos hospitales de la ciudad.

2. Clasificar los siguientes estudios en variables: Cualitativa nominal, cualitativa ordinal, cuantitativa discreta o cuantitativa continua.

A. Altura de una persona
B. El tiempo que emplea una persona en hacer una transacción en un banco
C. Comida favorita de un batallón

MEDIDAS DE POSICIÓN

Las medidas de posición dividen un conjunto de datos en grupos con el mismo número de individuos.

Para calcular las medidas de posición es necesario que los datos estén ordenados de menor a mayor.

CUARTILES

Los cuartiles son los tres valores de la variable que dividen a un conjunto de datos ordenados en cuatro partes iguales.

Q1, Q2 y Q3 determinan los valores correspondientes al 25%, al 50% y al 75% de los datos. Q2 coincide con la mediana.

1 Ordenamos los datos de menor a mayor.
2 Buscamos el lugar que ocupa cada cuartil mediante la expresión Cálculo de los cuartiles.

$$\frac{k \cdot N}{4},\ k = 1, 2, 3$$

Cálculo de los cuartiles para datos agrupados

$$\frac{k \cdot N}{4},\ k = 1, 2, 3$$

Buscamos el valor de en la tabla de las frecuencia acumuladas.

Luego para ese intervalo aplicamos la fórmula de los cuartiles.

$$Q_k = L_i + \frac{\frac{k \cdot N}{4} - F_{i-1}}{f_i} \cdot a_i \qquad k = 1, 2, 3$$

ACTIVIDAD 3

1. Calcula los cuartiles para la siguiente distribución de datos.

   1° 2° 3° 4° 5° 6° 7° 8° 9° 10° 11° 12° 13° 14° 15° 16° 17° 18° 19° 20°
   6 7 7 8 8 8 8 9 9 9 9 9 9 9 10 10 10 10 10 11

2. Calcula los cuartiles para la siguiente distribución de datos.

   10 11 11 12 12 12 12 12 13 13
   13 14 14 15 15 16 16 18 20

3. Los cincuenta alumnos que se presentaron al examen de Estadística en Junio obtuvieron las siguientes calificaciones. Determinas los cuartiles.

   7 3 2 4 5 1 8 6 1 53 2 4 9 8 1 0 2 4 12 5 6 5 4 7 1 3 0 5 8 6 3 4 0 10 2 5 7 4 0 2 1 5
   6 4 3 5 2 3

4. Calcula los cuartiles para los datos agrupados en las siguientes tablas.

a.

	f_i
[60, 63)	5
[63, 66)	18
[66, 69)	42
[69, 72)	27
[72, 75)	8
	100

b.

Estatura	(fi)
[145, 150)	5
[150, 155)	17
[155, 160)	7
[160, 165)	7
[165, 170)	4

c.

Intervalo de clase	Frecuencia
[70;78)	1
[78;86)	2
[86;94)	6
[94;102)	7
[102;110)	10
[110;118)	3
[118;126)	5
[126;134)	8
[134;142)	6
[142;150)	2

DECILES

Los deciles son los nueve valores que dividen la serie de datos en diez partes iguales. Los deciles dan los valores correspondientes al 10%, al 20%... y al 90% de los datos.

Cálculo de los deciles

$$\frac{k \cdot N}{10}, \quad k = 1, 2, ..9$$

Buscamos en la tabla de datos el intervalo donde se encuentre.

$$D_k = L_i + \frac{\frac{k \cdot N}{10} - F_{i-1}}{f_i} \cdot a_i \qquad k = 1, 2, ..9$$

ACTIVIDAD 4

Calcule los deciles para las siguientes frecuencias agrupadas.

a.

Estatura	(fi)
[145, 150)	5
[150, 155)	17
[155, 160)	7
[160, 165)	7
[165, 170)	4

b.

Intervalos de clase	fi
70 - 74	4
75 - 79	5
80 - 84	6
85 - 89	7
90 - 94	8

C.

Intervalos	Fi
16 – 21	8
22 – 27	12
28 – 33	18
34 – 39	17
40 – 45	17
46 – 51	12
52 – 57	8
58 – 63	9
64 69	3

PERCENTILES

Los percentiles son los 99 valores que dividen la serie de datos en 100 partes iguales. Los percentiles dan los valores correspondientes al 1%, al 2%... y al 99% de los datos.

Cálculo de los percentiles.

$$\frac{k \cdot N}{100}, \quad k = 1, 2, ..99$$

Buscamos el intervalo donde se encuentra , en la tabla de las frecuencias acumuladas.

$$P_k = L_i + \frac{\frac{k \cdot N}{100} - F_{i-1}}{f_i} \cdot a_i \qquad k = 1, 2, ..99$$

ACTIVIDAD 5

1. Para los datos agrupados de la actividad 4, calcula los percentiles 1, 10 , 30, 70, 99.
2. 37) Hallar los cuartiles, los deciles, P80 y P14 e interpretar su significado de los ejercicios.
   4a, 4b, 4c.
3. La tabla muestra una distribución de frecuencias de puntuaciones de un examen.

Puntuación	Número de estudiantes
[30 – 40)	1
[40 – 50)	3
[50 – 60)	11
[60 – 70)	21
[70 – 80)	43
[80 – 90)	32
[90 – 100)	9

Hallar:

2 Los cuartiles de la distribución.
3 P12 e interprete su significado.
4 P68 e interprete su significado.
5 La puntuación más baja alcanzada por el 25% más alto del curso y la más alta alcanzada por el 20% más bajo del curso.
6 ¿A qué porcentaje corresponde una calificación de 58 puntos?.

UNIDAD 3: MEDIDAS DE DISPERSIÓN

A pesar de la gran importancia de las medidas de tendencia central y de la cantidad de información que aportan individualmente, no hay que dejar de señalar que en muchas ocasiones esa información, no sólo no es completa, sino que puede inducir a errores en su interpretación. Veamos algunos ejemplos.

Consideremos dos grupos de personas extraídos como muestras respectivas de dos poblaciones distintas: el primero está compuesto por 100 personas que asisten a la proyección de una película para niños, y el segundo por 100 personas elegidas entre los

asistentes a una discoteca juvenil. Pudiera ocurrir que, aun siendo las distribuciones de las edades de ambos grupos muy distinta, la media y la mediana coincidieran para ambas. (Da un ejemplo concreto en que esto ocurra).

Las medidas de dispersión nos informan sobre cuánto se alejan del centro los valores de la distribución.

RANGO, AMPLITUD TOTAL O RECORRIDO

El rango se suele definir como la diferencia entre el mayor y el menor de los datos de una distribución estadística.
Es la medida de dispersión más sencilla y también, por tanto, la que proporciona menos información. Además, esta información puede ser errónea, pues el hecho de que no influyan más de dos valores del total de la serie puede provocar una deformación de la realidad.
Comparemos, por ejemplo, estas dos series:

Serie 1: 1 5 7 7 8 9 9 10 17
Serie 2: 2 4 6 8 10 12 14 16 18

Ambas series tienen rango 16, pero están desigualmente agrupadas, pues mientras la primera tiene una mayor concentración en el centro, la segunda se distribuye uniformemente a lo largo de todo el recorrido.

El uso de esta medida de dispersión, será pues, bastante restringido.

DESVIACIÓN MEDIA

En teoría, la desviación puede referirse a cada una de las medidas de tendencia central: media, mediana o moda; pero el interés se suele centrar en la medida de la desviación con respecto a la media, que llamaremos desviación media.

La desviación respecto a la media es la diferencia entre cada valor de la variable estadística y la media aritmética.

$$DM = \frac{\sum |x - \bar{x}|}{N}$$

Ejemplo: Calcular la desviación media de la distribución:

9, 3, 8, 8, 9, 8, 9, 18

$$\bar{X} = \frac{9+3+8+8+9+8+9+18}{8} = 9$$

$$D_{\bar{x}} = \frac{|9-9|+|3-9|+|8-9|+|8-9|+|9-9|+|8-9|+|9-9|+|18-9|}{8} = 2.25$$

DESVIACIÓN MEDIA PARA DATOS AGRUPADOS

Si los datos vienen agrupados en una tabla de frecuencias, la expresión de la desviación media es:

$$D_{\bar{x}} = \frac{|x_1 - \bar{x}|f_1 + |x_2 - \bar{x}|f_2 + \ldots + |x_n - \bar{x}|f_n}{N}$$

$$D_{\bar{x}} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|f_i}{N}$$

Calcular la desviación media de la distribución:

	x_i	f_i	$x_i \cdot f_i$	$\lvert x - \bar{x}\rvert$	$\lvert x - \bar{x}\rvert \cdot f_i$
[10, 15)	12.5	3	37.5	9.286	27.858
[15, 20)	17.5	5	87.5	4.286	21.43
[20, 25)	22.5	7	157.5	0.714	4.998
[25, 30)	27.5	4	110	5.714	22.856
[30, 35)	32.5	2	65	10.174	21.428
		21	457.5		98.57

ACTIVIDAD 6

1. Sea una distribución estadística que viene dada por la siguiente tabla:

x_i	61	64	67	70	73
f_i	5	18	42	27	8

Calcular el rango, desviación media

2. Calcular el rango y la desviación media para la siguiente serie de números: 5, 3, 6, 5, 4, 5, 2, 8, 6, 5, 4, 8, 3, 4, 5, 4, 8, 2, 5, 4.

3. Calcular la desviación media de una distribución estadística que viene dada por la siguiente tabla:

	[10, 15)	[15, 20)	[20, 25)	[25, 30)	[30, 35)
f_i	3	5	7	4	2

VARIANZA

La varianza es la media aritmética del cuadrado de las desviaciones respecto a la media de una distribución estadística.

La varianza se representa por σ^2.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N} \quad \sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{N}$$

Calcular la varianza de la distribución:
9, 3, 8, 8, 9, 8, 9, 18

$$\bar{x} = \frac{9+3+8+8+9+8+9+18}{8} = 9$$

$$\sigma^2 = \frac{(9-9)^2 + (3-9)^2 + (8-9)^2 + (8-9)^2 + (9-9)^2 + (8-9)^2 + (9-9)^2 + (18-9)^2}{8} = 15$$

VARIANZA PARA DATOS AGRUPADOS

$$\sigma^2 = \frac{x_1^2 f_1 + x_2^2 f_2 + \ldots + x_n^2 f_n}{N} - \bar{x}^2 \quad \sigma^2 = \sum_{i=1}^{n} \frac{x_i^2 f_i}{N} - \bar{x}^2$$

Calcular la varianza de la distribución de la tabla:

	x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[10,20)	15	1	15	225
[20,30)	25	8	200	5000
[30,40)	35	10	350	12250
[40,50)	45	9	405	18 225
[50, 60	55	8	440	24 200
[60,70)	65	4	260	16 900
[70,80)	75	2	150	11 250
		42	1820	88050

$$\bar{x} = \frac{1820}{42} = 43.33$$

$$\sigma^2 = \frac{88050}{42} - 43.33^2 = 218.94$$

PROPIEDADES DE LA VARIANZA

1. La varianza será siempre un valor positivo o cero, en el caso de que las puntuaciones sean iguales.

2. Si a todos los valores de la variable se les suma un número la varianza no varía.
3. Si todos los valores de la variable se multiplican por un número la varianza queda multiplicada por el cuadrado de dicho número.
4. Si tenemos varias distribuciones con la misma media y conocemos sus respectivas varianzas se puede calcular la varianza total.

Si todas las muestras tienen el mismo tamaño:

$$\sigma^2 = \frac{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2}{n}$$

Si las muestras tienen distinto tamaño:

$$\sigma^2 = \frac{k_1 \cdot \sigma_1^2 + k_2 \cdot \sigma_2^2 + \ldots + k_n \cdot \sigma_n^2}{k_1 + k_2 + \ldots + k_n}$$

OBSERVACIONES SOBRE LA VARIANZA

La varianza, al igual que la media, es un índice muy sensible a las puntuaciones extremas.
En los casos que no se pueda hallar la media tampoco será posible hallar la varianza.
La varianza no viene expresada en las mismas unidades que los datos, ya que las desviaciones están elevadas al cuadrado.

DESVIACIÓN ESTANDAR

La desviación típica es la raíz cuadrada de la varianza.
Es decir, la raíz cuadrada de la media de los cuadrados de las puntuaciones de desviación.
La desviación típica se representa por σ.

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{N}} \quad \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}}$$

Calcular la desviación típica de la distribución:

9, 3, 8, 8, 9, 8, 9, 18

$$\bar{x} = \frac{9+3+8+8+9+8+9+18}{8} = 9$$

$$\sigma^2 = \sqrt{\frac{(9-9)^2 + (3-9)^2 + (8-9)^2 + (8-9)^2 + (9-9)^2 + (8-9)^2 + (9-9)^2 + (18-9)^2}{8}} = 3.87$$

DESVIACIÓN TÍPICA PARA DATOS AGRUPADOS

$$\sigma = \sqrt{\frac{x_1^2 f + x_2^2 f_2 + \ldots + x_n^2 f_n}{N} - \bar{x}^2} \quad \sigma = \sqrt{\sum_{i=1}^{n} \frac{x_i^2 f_i}{N} - \bar{x}^2}$$

Calcular la desviación típica de la distribución de la tabla:

	x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[10, 20)	15	1	15	225
[20, 30)	25	8	200	5000
[30,40)	35	10	350	12 250
[40, 50)	45	9	405	18 225
[50, 60)	55	8	440	24 200
[60,70)	65	4	260	16 900
[70, 80)	75	2	150	11 250
		42	1 820	88 050

$$\bar{x} = \frac{1820}{42} = 43.33$$

$$\sigma = \sqrt{\frac{88050}{42} - 43.33^2} = 14.797$$

PROPIEDADES DE LA DESVIACIÓN TÍPICA

1. La desviación típica será siempre un valor positivo o cero, en el caso de que las puntuaciones sean iguales.
2. Si a todos los valores de la variable se les suma un número la desviación típica no varía.
3. Si todos los valores de la variable se multiplican por un número la desviación típica queda multiplicada por dicho número.
4. Si tenemos varias distribuciones con la misma media y conocemos sus respectivas desviaciones típicas se puede calcular la desviación típica total.

Si todas las muestras tienen el mismo tamaño:

$$\sigma = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2}{n}}$$

Si las muestras tienen distinto tamaño:

$$\sigma = \sqrt{\frac{k_1 \cdot \sigma_1^2 + k_2 \cdot \sigma_2^2 + \ldots + k_n \cdot \sigma_n^2}{k_1 + k_2 + \ldots + k_n}}$$

ACTIVIDAD 7

1. Hallar la varianza y la desviación típica de la siguiente serie de datos:

12, 6, 7, 3, 15, 10, 18, 5.

2. Un pediatra obtuvo la siguiente tabla sobre los meses de edad de 50 niños de su consulta en el momento de andar por primera vez:

Meses	Niños
9	1
10	4
11	9
12	16
13	11
14	8
15	1

a. Calcular la varianza y la desviación típica.

3. Calcular la desviación típica de una distribución estadística que viene dada por la siguiente tabla:

Xi	[10,15)	[15,20)	[20,25)	[25,30)	[30,35)
f_i	3	5	7	4	2

4. Las alturas de los jugadores de un equipo de baloncesto vienen dadas por la tabla:

Altura	Nº de jugadores
[170, 175)	1
[175, 180)	3
[180, 185)	4
[185, 190)	8
[190, 195)	5
[195, 2.00)	2

Calcular la varianza y la desviación típica.

UNIDAD 4: REPRESENTACIONES GRAFICAS

Una representación gráfica o grafico estadístico es un resumen visual de la tabla de frecuencias, que nos permite la representación en forma más resumida y total del experimento o fenómeno en estudio.

Las representaciones gráficas son muy útiles como apoyos e incluso sustitutos de las tablas de frecuencia y como una herramienta para el análisis de los datos, lo que los convierte en el medio más efectivo para la presentación, descripción, resumen y análisis de la información.

Para el caso de variables cualitativas se suelen usar dos clases de representaciones graficas que son:
Diagramas de barras
Diagrama circular

Diagrama de barras

El diagrama de barras (o gráfico de barras) es la representación gráfica, en ejes cartesianos, de la información dada en una tabla de frecuencias, este es construido por barras rectangulares cuya altura es proporcional a la frecuencia de cada uno de los valores de la variable.

Las principales características del diagrama de barras son:

- En el eje x (**eje de abscisas)** se colocan las clases de la variable cualitativa.
- En el eje y (**eje de ordenadas) se** colocan las barras proporcionales a la frecuencia absoluta o frecuencia relativa de la clase.
- Todas las barras deben construirse separadas y deben tener el **mismo ancho.**

Ejemplo1

Los resultados obtenidos de preguntar a un grupo de alumnos del grado 11 de la I.E. Juan Manuel González, por la asignatura de su preferencia, fueron resumidos en la siguiente tabla de frecuencias.

Asignatura de preferencia	Frecuencia f	Frecuencia relativa f_r	Porcentaje %
Ética y valores	5	5/50	10
Emprendimiento	9	9/50	18
Ingles	10	10/50	20
Matemáticas	9	9/50	18
Química	6	6/50	12
Sociales	4	4/50	8
Tecnología	7	7/50	14
Total	50	1	100

Para visualizar mejor los datos, construir un **diagrama de barras verticales con la frecuencia absoluta**

Solución

1. Identificamos la variable y las clases que la componen.

Variable: Asignatura de preferencia

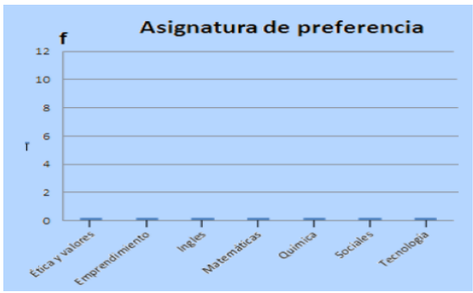
Clases de la Variable:

· Ética y valores
· Emprendimiento
· Ingles
· Matemáticas
· Química
· Sociales
· Tecnología

2. Identificamos cual es la frecuencia que deseamos graficar. Para este caso es la frecuencia absoluta la cual tiene un valor máximo de 10 que corresponde a la frecuencia de inglés.

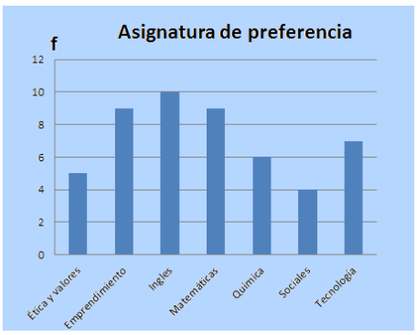
3. Construimos un plano cartesiano con una escala adecuada, lo que quiere decir que en eje x debemos seleccionar 7 divisiones igualmente espaciadas que servirán para ubicar las

clases y en eje y, tenemos que el valor máximo a representar es 10, con este dato procedemos a escalar el eje y; para nuestros caso usaremos divisiones de 2 en 2.



4. iniciamos a graficar los datos de la tabla de frecuencia. Por ejemplo, sabemos que Ética y valores tienen una frecuencia de 5, para representar esta situación en el diagrama de barras nos ubicamos en el eje x en Ética y valores y construimos una barra rectangular de altura igual a la frecuencia la cual se ubica en el eje y.

5. realizando el mismo proceso del paso 4 para las demás clases obtenemos el diagrama de barras



ACTIVIDAD 8

Se preguntó a un total de 1000 residentes de Minnesota que estación del año preferían. Los resultados se encuentran en la siguiente tabla de frecuencias. Realizar gráfica.

Estación del año	Frecuencia f
Invierno	100
Otoño	200
Verano	400
Primavera	300
Total	1000

Diagrama circular

El **diagrama circular** (también llamado **diagrama de sectores o diagrama de pastel**) es la representación gráfica en sectores circulares, de la información dada en una tabla de frecuencias.

Para elaborar este tipo de gráficos, se debe dividir el círculo en partes proporcionales a los valores de las frecuencias dadas.

Cálculo de los sectores

Un círculo completo son 360 °. El sector que ocupa cada clase por definición es proporcional a la frecuencia de esta. Por ejemplo, si una clase tiene una frecuencia relativa porcentual del 50%, su sector del círculo tendrá un ángulo de 180°.

Lo cual puede ser comprobado con una regla de tres simple directa así:

Si una frecuencia relativa porcentual del 100 % equivale a 360 ° ¿A cuántos grados equivale una frecuencia relativa porcentual del 50 %?

$$\begin{array}{ccc} 100\,\% & \longrightarrow & 360\,^\circ \\ 50\,\% & \longrightarrow & x \end{array}$$

$$x = \frac{360\,^\circ \cdot 50\,\%}{100\,\%} = 180\,^\circ$$

Respuesta: una frecuencia relativa porcentual de 50 % equivale 180°

Ejemplo

Realizar el diagrama circular de la siguiente tabla de frecuencias

Estación del año	Frecuencia f
Invierno	100
Otoño	200
Verano	400
Primavera	300
Total	1000

Solución

1. Para el desarrollo del problema lo primero que debemos hacer es agregar una columna a la tabla de frecuencias, en la cual se consignaran los grados de los sectores circulares que deben ocupar los datos consignados en la tabla.

Estación del año	Frecuencia f	Grados °
Invierno	100	
Otoño	200	
Verano	400	

Primavera	300	
Total	1000	

2. Calculamos los grados del sector circular que le corresponde a cada clase, aplicando una regla de tres directa, como se explica a continuación:
   Recuerde que el total de los datos ocupan 360 °
   Si una frecuencia 1000 equivale a 360 ° ¿A cuántos grados equivale una frecuencia de 100?

$$\begin{array}{rcl} 1000 & \longrightarrow & 360\,° \\ 100 & \longrightarrow & x \end{array}$$

$$x_{invierno} = \frac{360\,° * 100}{1000} = 36$$

Respuesta: una frecuencia de 100 de 50 % equivale a 36 °
Se realiza el mismo procedimiento para el resto de las clases.

$$\begin{array}{rcl} 1000 & \longrightarrow & 360\,° \\ 400 & \longrightarrow & x \end{array}$$

$$x_{invierno} = \frac{360\,° * 400}{1000} = 108$$

$$\begin{array}{rcl} 1000 & \longrightarrow & 360\,° \\ 400 & \longrightarrow & x \end{array}$$

$$x_{invierno} = \frac{360\,° * 300}{1000} = 144$$

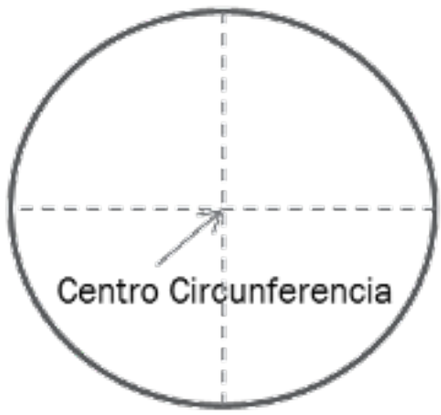
$$\begin{array}{rcl} 1000 & \longrightarrow & 360\,° \\ 200 & \longrightarrow & x \end{array}$$

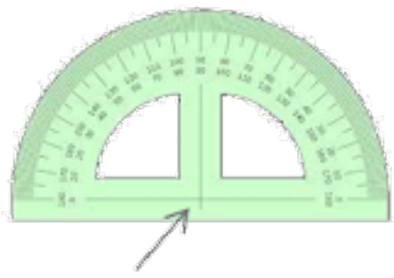
$$x_{invierno} = \frac{360\,° * 200}{1000} = 72$$

3. Consignamos los valores obtenidos en la tabla de frecuencias.

Estación del año	Frecuencia f	Grados °
Invierno	100	36
Otoño	200	72
Verano	400	144
Primavera	300	108
Total	1000	360

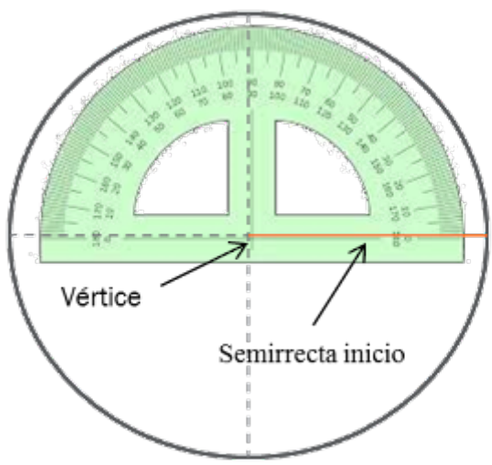
4. Realice el trazo de la circunferencia que utilizara como "pastel" y trace sobre esta con líneas punteadas una línea vertical y otra horizontal que pasen por el centro.
Tenga a la mano un transportador.





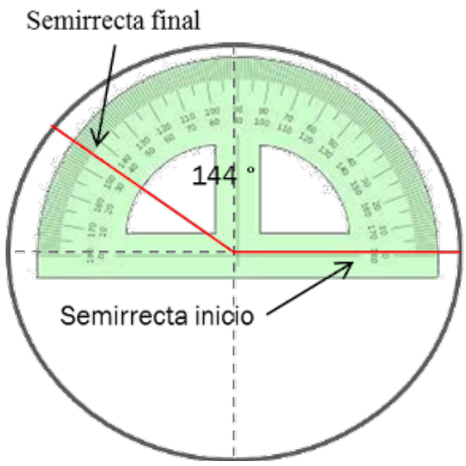
5. Para construir el ángulo que ocupa un sector circular se debe tomar como vértice (punto donde se encuentran dos o más semirrectas o segmentos que conforman un ángulo) de este el centro del diagrama de pastel.
Para construir un ángulo de un sector circular del diagrama de pastel siga el siguiente procedimiento:

- Trace la semirrecta de inicio desde el vértice (centro de la circunferencia).
- Ubique el transportador de forma tal que coincidan el centro del transportador y la circunferencia (con esto garantizamos que el vértice es el centro de la circunferencia).
- Alinee la semirrecta de inicio con la marca de cero grados del transportador
- Marque el ángulo que desea representar y trace la semirrecta final.

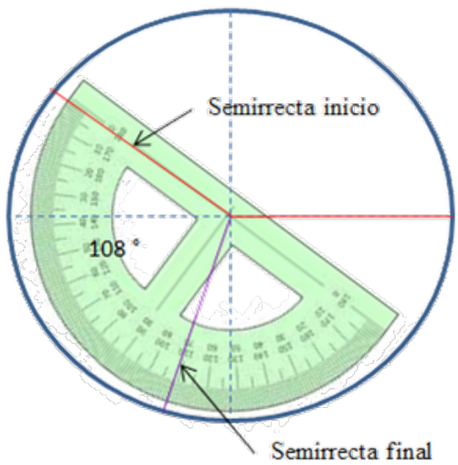


6. Con el transportador ubicado en el vértice y alineado con la semirrecta de inicio se miden 144 ° que equivale al ángulo del sector circular de la clase verano, y se traza la

semirrecta de final. Quedando de esta forma delimitado el sector circular que representa la clase verano entre la semirrecta de inicio y la semirrecta final.

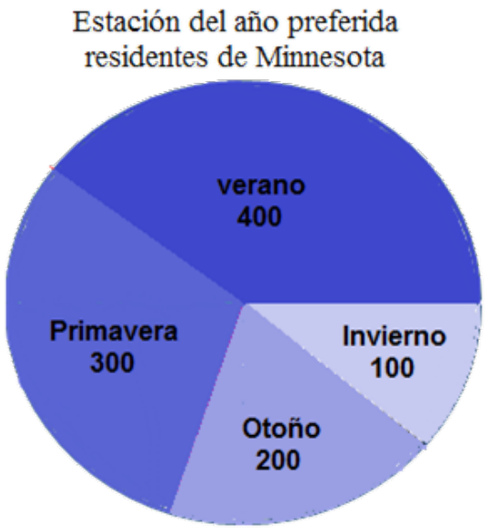


7. Gire el transportador hasta que quede la línea de 0 ° alineada con la semirrecta final que acaba de construir, esta se convertirá en la semirrecta de inicio para medir el nuevo ángulo. Siga el procedimiento del paso 5 para medir el ángulo de la clase primavera. Ver la figura que aparece a continuación.



8. Repita el procedimiento anterior para las clases sucesivas hasta que divida el diagrama circular en la cantidad de clases que aparecen en la tabla de frecuencias. Ver la figura que aparece a continuación.

9. Finalmente coloree los sectores circulares de un color diferente y de ser posible se ubica sobre estos el nombre de las clases y la frecuencia.
Se le asigna un titulo que ponga de manifiesto la variable que se esta representando en el diagrama circular.



ACTIVIDAD 9

1. Una empresa del municipio de Dosquebradas desea saber el estado civil de sus empleados, para lograrlo le aplico una encuesta a 400. La información es tabulada de la siguiente manera:

Estado Civil	*f*
Soltero	75
Casado	200
Viudo	50
Separado	75
Total	400

a. Determinar Población, muestra y variable (Tipo, Clases).
b. Completar la tabla de frecuencias (frecuencia relativa y frecuencia relativa porcentual)
c. Realizar la representación gráfica de la información con un diagrama de barras y un diagrama circular.

UNIDAD 5: DISTRIBUCION DE PROBABILIDAD DISCRETA

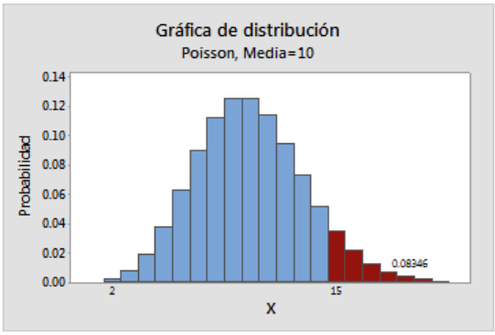
Una **distribución discreta** describe la **probabilidad** de ocurrencia de cada valor de una variable aleatoria **discreta**. ... Con una **distribución de probabilidad discreta**, cada valor posible de la variable aleatoria **discreta** puede estar asociado con una **probabilidad** distinta de cero.

Ejemplo:

Con una distribución discreta, a diferencia de una distribución continua, usted puede calcular la probabilidad de que X sea exactamente igual a algún valor. Por ejemplo, puede utilizar la distribución discreta de Poisson para describir el número de quejas de clientes en un día. Supongamos que el número promedio de quejas por día es 10 y usted desea saber la probabilidad de recibir 5, 10 y 15 quejas de clientes en un día.

x	P (X = x)
5	0.037833
10	0.12511
15	0.034718

Usted también puede visualizar una distribución discreta en una gráfica de distribución para ver las probabilidades entre los rangos.



Gráfica de distribución del número de quejas de clientes

Las barras sombreadas en este ejemplo representan el número de ocurrencias cuando las quejas diarias de los clientes son 15 o más. La altura de las barras suma 0.08346; por lo tanto, la probabilidad de que el número de llamadas por día sea 15 o más es 8.35%.

Distribución Binomial

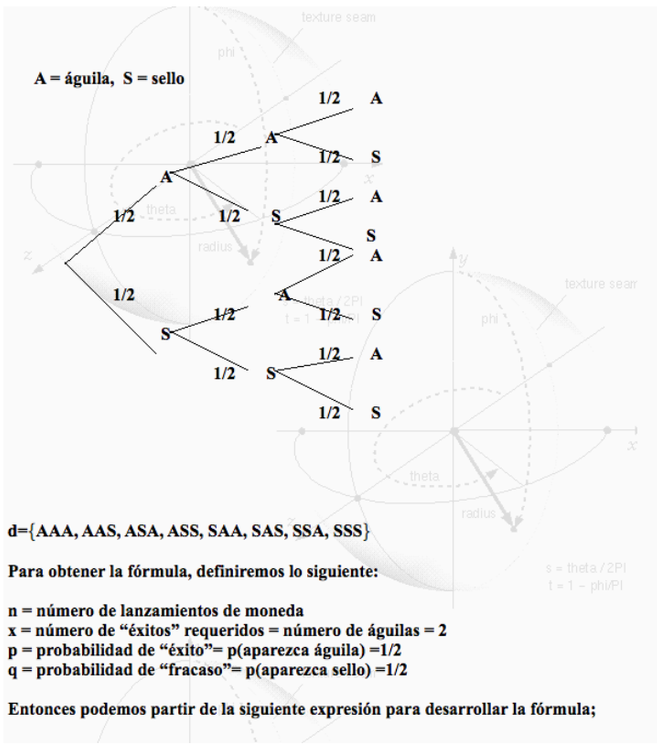
Las características de esta distribución son:

a) En los experimentos que tienen este tipo de distribución, siempre se esperan dos tipos de resultados, ejemplo. Defectuoso, no defectuoso, pasa, no pasa, etc, etc., denominados arbitrariamente “éxito” (que es lo que se espera que ocurra) o “fracaso” (lo contrario del éxito).
b) Las probabilidades asociadas a cada uno de estos resultados son constantes, es decir no cambian.
c) Cada uno de los ensayos o repeticiones del experimento son independientes entre sí.
d) El número de ensayos o repeticiones del experimento (n) es constante.

A partir de un ejemplo. Desarrollaremos una fórmula que nos permita cualquier problema que tenga este tipo de distribución.
Ejemplo:
Se lanza al aire una moneda normal 3 veces, determine la probabilidad de que aparezcan 2 águilas.

Solución:
Antes de empezar a resolver este problema, lo primero que hay que hacer es identificarlo como un problema que tiene una distribución binomial, y podemos decir que efectivamente así es, ya que se trata de un experimento en donde solo se pueden esperar dos tipos de resultados al lanzar la moneda, águila o sello, cutas probabilidades de ocurrencia son constantes, cada uno de los lanzamientos es independiente de los demás y el número de ensayos o repeticiones del experimento son constantes, n = 3.

Para dar solución a este problema, lo primero que hay que hacer es un diagrama de árbol, en donde representaremos los tres lanzamientos, de ahí se obtendrá el espacio muestral y posteriormente la probabilidad pedida, usando la fórmula correspondiente.



Entonces podemos partir de la siguiente expresión para desarrollar la fórmula;

P (aparezcan 2 águilas) = (No. De ramas del árbol en donde ap. 2 águilas) (probabilidad asociada a cada rama)

Entonces el número de ramas en donde aparecen dos águilas se puede obtener;

Enumerando las ramas de interés, estas serían: AAS, ASA, SAA, ¿QUÉ TIPO DE ARREGLOS SON ESTOS ELEMENTOS DEL ESPACIO MUESTRAL?, Son permutaciones en donde algunos objetos son iguales, entonces, el número de ramas se puede obtener con la fórmula correspondiente,

$$nP_{x1,x2,\ldots xk} = \frac{n!}{x_1!\,x_2!\ldots x_k!}$$

donde $n = x_1+x_2+\ldots+x_k$

sustituyendo en esta fórmula, tenemos lo siguiente;

$$nP_{x,n-x} = \frac{n!}{x!(n-x)!}$$

esta fórmula puede ser sustituida por la de combinaciones, solo en el caso de dos tipos de objetos, si hay más de dos tipos de objetos, definitivamente solo se usa la fórmula original, como se observará en el caso de la distribución multinomial, pero ¿porqué vamos a cambiar de fórmula?, simplemente porque en todos los libros de texto que te encuentres vas a encontrar la fórmula de combinaciones en lugar de la de permutaciones, que es la siguiente,

$$nCx = \frac{n!}{x!(n-x)!}$$

y sustituyendo valores, nos damos cuenta de que efectivamente son 3 las ramas de interés, que son donde aparecen dos águilas, donde n = 3, x = 2.

$$_3C_2 = \frac{3!}{2!(3-2)!} = \frac{3!}{2!1!} = \frac{3x2!}{2!1!} = 3ramas$$

¿Y la probabilidad asociada a cada rama?
Probabilidad asociada a cada rama = p(águila)*p(águila)*p(sello)= p*p*q = p^2q=

$$= p^x q^{n-x}$$

Luego la fórmula de la distribución Binomial sería:

$$p(n,x,p) = {}_nC_{n-x}p^x q^{n-x}$$

donde:
p(*x, n, p*) = probabilidad de obtener en *n* ensayos *x* éxitos, cuando la probabilidad de éxito es *p*

Dando solución al problema de ejemplo tenemos lo siguiente:
n = 3, x = 2, p = ½

$$p(n=3,x=2,p=1/2) = {}_3C_2(1/2)^2(1/2)^{3-2} = \frac{3!}{2!1!} * \frac{1}{4} * \frac{1}{2} = 3 * \frac{1}{8} = \frac{3}{8}$$

Para calcular la media y la desviación estándar de un experimento que tenga una distribución Binomial usaremos las siguientes fórmulas:

Media o valor esperado.

$$\mu = nP$$

Donde:
n = número de ensayos o repeticiones del experimento
P = probabilidad de éxito o la probabilidad referente al evento del cual se desea calcular la media que se refiere la media
Q = complemento de P

Desviación estándar.

$$\sigma = \sqrt{nPQ}$$

Ejemplos:

1. Se dice que el 75% de los accidentes de una planta se atribuyen a errores humanos. Si en un período de tiempo dado, se suscitan 5 accidentes, determine la probabilidad de que; a) dos de los accidentes se atribuyan a errores humanos, b) como máximo 1 de los accidentes se atribuya a errores de tipo humano, c) tres de los accidentes no se atribuyan a errores humanos.

Solución:

a) n = 5

x = variable que nos define el número de accidentes debidos a errores humanos

x = 0, 1, 2, 5 accidentes debidos a errores de tipo humano

p = p(éxito) = p (un accidente se deba a errores humanos) = 0.75

q = p(fracaso) = p (un accidente no se deba a errores humanos) = 1-p = 0.25

$$p(x=2,n=5,p=0.75)={}_5C_2(0.75)^2(0.25)^{5-2}=(10)(0.5625)(0.015625)=0.08789$$

b) $$p(x=0,1,n=5,p=0.75)=p(x=0)+p(x=1)={}_5C_0(0.75)^0(0.25)^{5-0}+$$

$${}_5C_1(0.75)^1(0.25)^{5-1}=0.000976+0.014648=0.015624$$

c) En este caso cambiaremos el valor de p;

n =5

x = variable que nos define el número de accidentes que no se deben a errores de tipo humano

x = 0, 1, 2, 5 accidentes debidos a errores humanos

p = p (probabilidad de que un accidente no se deba a errores humanos) = 0.25

q = p (probabilidad de que un accidente se deba a errores humanos) = 1-p = 0.75

$$p(x=3,n=5,p=0.25)={}_5C_3(0.25)^3(0.75)^{5-3}=(10)(0.015625)(0.5625)=0.08789$$

2. Si la probabilidad de que el vapor se condense en un tubo de aluminio de cubierta delgada a 10 atm de presión es de 0.40, si se prueban 12 tubos de ese tipo y bajo esas condiciones, determine la probabilidad de que: a) el vapor se condense en 4 de los tubos, b) en más de 2 tubos se condense el vapor, c) el vapor se condense en exactamente 5 tubos.

Solución:

a) n =12

x = variable que nos define el número de tubos en que el vapor se condensa

x = 0, 1, 2, 3,12 tubos en el que el vapor se condensa

p =p (se condense el vapor en un tubo de Al a 10 atm) = 0.40

q = p (no se condense el vapor en un tubo de Al a 10 atm) = 1-p=0.60

$$p(x=4,n=12,p=0.40)={}_{12}C_4(0.40)^4(0.60)^{12-4}=(495)(0.0256)(0.016796)=$$

$$=0.21284$$

b) p(X=3, 4, ...,12, n=12, p=0.40) = p(x=3)+p(x=4)+…+p(x=12)= 1-[p(x=0,1,2)]=

$$= 1 - \left[{}_{12}C_0 (0.40)^0 (0.60)^{12-0} + {}_{12}C_1 (0.40)^1 (0.60)^{12-1} + {}_{12}C_2 (0.40)^2 (0.60)^{12-2} \right]$$

$$= 1 - [0.002176 + (12)(0.4)(0.003627) + (66)(0.16)(0.006047)]$$

= 1- [0.002176+0.0174096+0.06385632]= 1- 0.08344192= 0.91656

c)

$$p(x = 5, n = 12, p = 0.40) = {}_{12}C_5 (0.40)^5 (0.6)^{12-5} = (792)(0.01024)(0.0279936) =$$

= 0.22703

3. La probabilidad de que el nivel de ruido de un amplificador de banda ancha exceda de 2 dB (decibeles) es de 0.15, si se prueban 10 amplificadores de banda ancha, determine la probabilidad de que; a) en solo 5 de los amplificadores el nivel de ruido exceda los 2 dB, b) por lo menos en 2 de los amplificadores, el ruido exceda de 2 dB, c)que entre 4 y 6 amplificadores no se excedan de los 2 dB, d)encuentre el número esperado de amplificadores que se exceden de un nivel de ruido de 2dB y su desviación estándar.

Solución:
a) n =10
x =variable que nos define el número de amplificadores de banda ancha que su nivel de ruido excede de 2 dB
x = 0, 1, 2, ...,10 amplificadores en los que el nivel de ruido excede de los 2 dB
p = P (un amplificador exceda su nivel de ruido de 2 dB) = 0.15
q = p (un amplificador no exceda su nivel de ruido de 2 dB =1-p= 0.85

$$p(x = 5, n = 10, p = 0.15) = {}_{10}C_5 (0.15)^5 (0.85)^{10-5} = (252)(0.00007593)(0.4437053) =$$

= 0.00849

b) p (x=2, 3...,10, n=10, p=0.15) = 1- p (x = 0,1) =

$$= 1 - \left[{}_{10}C_0 (0.15)^0 (0.85)^{10-0} + {}_{10}C_1 (0.15)^1 (0.85)^{10-1} \right]$$

= 1 – [(0.19687+(10) (0.15) (0.231617)]=1-0.544296 = 0.455705

c) n=10
x= variable que nos define el número de amplificadores de banda ancha que su nivel de ruido no excede de 2 dB
x= 0, 1, 2...,10 amplificadores que su nivel de ruido no excede de los 2 dB
p = p (un amplificador no exceda su nivel de ruido de 2 dB) = 0.85
q = p (un amplificador exceda su nivel de ruido de 2 dB) = 1- p = 0.15

$$p(x = 4,5,6, n = 10, p = 0.085) = {}_{10}C_4 (0.85)^4 (0.15)^{10-4} + {}_{10}C_5 (0.85)^5 (0.15)^{10-5} + {}_{10}C_6 (0.85)^6 (0.15)^{10-6} =$$

= (210) (0.522)(0.00001139)+(252)(0.4437)(0.000075937)+(210)(0.3771495)(0.00005063)=
=0.001249 + 0.00849 + 0.00400997 = 0.01374897

d) n=10, p=0.15, q=1-p=0.85

$$\mu = np = (10)(0.15) = 1.5 \cong 2 amplificadores$$

Interpretación:
Se espera que 2 de los 10 amplificadores probados se excedan de un nivel de ruido de 2 Db

$$\sigma = \sqrt{npq} = \sqrt{(10)(0.15)(0.85)} = 1.1291 \cong \pm 1 amplificador$$

Interpretación:
Este experimento puede variar en 2 ± 1 amplificador, esto es, de 1 a 3 amplificadores que se excedan de un nivel de ruido de 2 dB

ACTIVIDAD 10

La probabilidad de que a un cliente nuevo le guste la matehamburguesa de Jorge es de 0,8. Si llegan 5 clientes nuevos a la cafetería, ¿cuál es la probabilidad de que solo a 3 de ellos les guste la matehamburguesa?

ACTIVIDAD 11

De todas las flores plantadas por una empresa de jardinería, el 90% sobrevive. Si se plantan 10 flores ¿cuál es la probabilidad de que 9 o más sobrevivan?

ACTIVIDAD 12

Considere un experimento binomial con dos ensayos y $p=0,4$.
a) Calcular la probabilidad de no obtener ningún éxito.
b) Calcular la probabilidad de obtener al menos 1 éxito.

ACTIVIDAD 13

Considere un experimento binomial con 10 ensayos y p = 0,9.
a) Calcular la probabilidad de obtener 9 éxitos.
b) Calcular la probabilidad de obtener 9 o más éxitos.

ACTIVIDAD 13

En una tienda, el 75 % de todas las compras se hacen con tarjeta de crédito. Sea X = el número entre diez compras seleccionadas al azar realizadas con tarjetas de crédito. Entonces X~B (10; 0,75).
a) Encontrar la función de probabilidad de X.
b) Calcular el valor esperado de X.
c) Calcule la varianza de X.
d) Calcule la desviación estándar de X.

4. GLOSARIO:

Moda: El valor mas alto de los registros.
Media: El valor promedio de los datos, es decir, todos los datos estarán cerca a la media.
Clase: También se conoce como intervalo, es la agrupación de una cantidad de valores.
Clase modal: intervalo en donde se encuentra la frecuencia absoluta mas alta.
Límite inferior: Es el extremo inferior de una clase o intervalo.
Límite superior: Es el extremo superior de una clase o intervalo.
Abscisa: El eje x's del plano cartesiano.
Ordenada: El eje y del plano cartesiano.
Variable discreta: Es la variable que puede asumir un número finito o contable de valores.
Variable continua: Es la variable que puede asumir un número infinito o incontable de valores.

Cuartiles: Valores de la variable que dividen a un conjunto de datos ordenados en 4 partes iguales.
Deciles: Valores de la variable que dividen a un conjunto de datos ordenados en 10 partes iguales.
Percentil: Es un valor de los 99 valores que dividen un conjunto ordenado en 100 partes iguales.
Varianza: Es una medida de dispersión de los datos, que representa la variabilidad o la fluctuación de todos los datos de una muestra con respecto a la media.

5. REFERENTES BILBIOGRAFÍCOS:

[1] *Obregón* Sanin, I. (1975). *Teoría de la probabilidad*.
[2] H. Runyon. (1973). *Estadística general.* Fondo educativo interamericano.

6. CONTROL DEL DOCUMENTO:

Autor (es)	Nombre	Cargo	Dependencia	Fecha
	John Edisson Tunubalá Morales	Docente	Matemáticas	Abril de 2020

7. CONTROL DE CAMBIOS: (diligenciar únicamente si realiza ajustes a la guía)

Autor (es)	Nombre	Cargo	Dependencia	Fecha	Razón del cambio