TEXTO GUÍA DE ESTADÍSTICA I

YURLYS LAUDITH ESCALANTE SALAS
LIZ ANGÉLICA HOLLMAN VERGARA

UNIVESIDAD TECNOLÓGICA DE BOLÍVAR
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
CARTAGENA
2010

TEXTO GUÍA DE ESTADÍSTICA I

YURLYS LAUDITH ESCALANTE SALAS
LIZ ANGÉLICA HOLLMAN VERGARA

Proyecto de grado presentado como requisito para optar por el título de ingeniería Industrial

ROBERTO E. GÓMEZ FERNÁNDEZ
INGENIERO INDUSTRIAL
Especialista en Estadística Aplicada
Especialista en Finanzas

UNIVESIDAD TECNOLÓGICA DE BOLÍVAR
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
CARTAGENA
2010

Nota de aceptación

____

____

____

____

____

____

____
Firma del presidente del jurado

____
Firma del jurado

____
Firma del jurado

Cartagena, 24 de Octubre de 2010

DEDICATORIA

Dedico este gran logro a Dios por darme las fuerzas para luchar y mejorar cada día para alcanzar las metas que me propongo a nivel personal y profesional.
A mi mamá Edith Salas por su amor, perseverancia y la educación que me ha brindado en todos estos años. A mi papá José Luis por enseñarme a ser una persona objetiva y crítica en esta vida. A mis hermanos Yesica, José Luis y Edith Carolina por su sencillez y el ánimo que me brindaron en cada uno de los momentos difíciles.
A mi tutor de tesis por su colaboración. A mis amigos especialmente Liz Angélica, María Alejandra, Tere, Antonio, Camilo, Juan Carlos, Keilly, Stephanie, Betin, Juan Diego, Angélica y a todas aquellas personas que de una u otra forma colaboraron en el transcurso de esta maravillosa carrera.

Yurlys Laudith Escalante Salas

Dedico este logro a Dios por Ayudarme siempre en Todo momento de mi vida,
A mi Mamá por su amor, paciencia y lucha incondicional para sacarnos adelante, mi Mami Linda por su cariño y colaboración en el transcurso de mi carrera, a mi Papá que se que en todo momento me esta observando y intercediendo por mi desde el cielo, a mi hermanita por su cariño, a toda mi familia que estuvieron siempre pendientes de que cumpliera este logro, a mi novio, Fabio Andrés, a mi compañera de tesis, Yurlys, amigos especialmente a Yira, Juan, Luz, Jesica, Héctor y Angélica; que de alguna u otra forma colaboraron con mi crecimiento personal y profesional en todos estos años.
Dedico este trabajo muy especialmente a Mi porque he luchado para que se cumpla.

Liz Angélica Hollman Vergara

AGRADECIMIENTOS

Gracias a Dios por darnos la perseverancia necesaria para llevar a cabo este proyecto. A nuestros padres y familiares por apoyarnos económicamente en el desarrollo de éste logro.

Gracias a nuestro Director de proyecto, Roberto E. Gómez Fernández, por ser nuestro guía y poner al servicio sus conocimientos para el desarrollo de este proyecto.

A Yira Laguna Morantes, Gerente de la planta de concretos de la empresa Laguna Morantes S.A., por brindarnos información con el fin de plantar casos aplicados a empresas de Cartagena.

A Fernando Figueroa, Auditor Interno de Gyptec S.A., y Francisco Solano, Exgerente Administrativo de Gyptec S.A., por permitirnos conocer los procesos de la empresa, de tal manera éstos sirvieran como guía para el planteamiento de ejemplos y ejercicios prácticos.

A Daniel Toscano, Ingeniero Civil de la UTB, por ayudarnos en el desarrollo de algunas demostraciones contenidas en el texto.

A Alicia E. Bozzi Martínez, Coordinadora de Proyectos del Proyecto Cartagena Como Vamos, por suministrarnos información útil para la proposición de casos.

Finalmente, les extendemos nuestro más sincero agradecimiento a todas aquellas personas y profesionales que colaboraron de una u otra manera en el proyecto.

Yurlys Laudith Escalante Salas

Liz Angélica Hollman Vergara

CONTENIDO

INTRODUCCION

La estadística y la probabilidad esta presente en muchos contextos de la vida, desde conocer el tipo de sexo de un bebe hasta llegar a predecir que números pueden salir en un juego de azar, permitiendo analizar información, tomar decisiones y realizar conclusiones a partir de inferencias estadísticas.

La estadística es una herramienta muy importante en cuanto proporciona diferentes métodos como la predicción, recolección de información, descripción gráficos; para la solución de problemas, la descripción y análisis de procesos, el mejoramiento continuo de los mismos, el diseño y desarrollo de nuevos productos, procesos y sistemas.

En este texto, se aborda conceptos básicos matemáticos para facilitar la comprensión de la estadística, los métodos para resumir y describir datos. Además de la teoría de probabilidad, variables aleatorias continuas y discretas, distribuciones de probabilidad para variables aleatorias continuas y discretas, con el fin de incluir toda la temática de los planes de estudio del curso de Estadística I~~.~~

Es importante destacar, que en este texto guía se plantean casos prácticos en los cuales se deben aplicar los métodos estadísticos y el uso de herramientas tecnológicas, tales como el Excel y Sthapgraphic para la solución de problemas de la cotidianidad.

La estadística es aplicada en diferentes áreas como la investigación científica, economía, psicología, sociología, ingeniería y otras. Este texto ha sido enfocado básicamente para las carreras de Ingenierías, Ciencias Sociales y política, Psicología, Administración y Negocios, con el propósito que el estudiante se familiarice con la utilización de técnicas estadísticas.

Capitulo 1

Fundamentos de Matemáticas.

INTRODUCCION

La estadística se basa en principios y reglas matemáticas usadas como herramientas para el análisis e interpretación de la información contenida en un conjunto de datos, cuyo análisis e interpretación facilitan el entendimiento de comportamientos de diferentes componentes en distintas situaciones.

El análisis matemático a lo largo de la existencia ha colaborado en la búsqueda de soluciones a grandes problemas como fueron el cálculo de la recta tangencial ó problemas de derivación, este tipo de análisis ayuda en cualquier problema en las ciencias ya que utiliza la lógica en su respuesta.

En este capítulo se presenta un tratado básico que contiene los conceptos fundamentales y métodos matemáticos, para formar una visión general de su importancia en la estadística y en el desarrollo de la misma, a través de razonamientos lógicos plasmados en ejercicios resueltos que muestran la interacción de la matemática y la estadística, además de herramientas informáticas como software –Statgraphics-, hojas de cálculo –Excel- y calculadora.

COMPETENCIAS:

Reconocimeinto de la relación entre un conjunto de datos y su representación.

Utilizacion de números reales en sus diferentes representaciones y en diversos contextos.

Resolucion de problemas y simplifico cálculos usando propiedades y relaciones de los números reales y de las relaciones y operaciones entre ellos.

Utilizacion de la notación científica para representar medidas de cantidades de diferentes magnitudes.

Identificacion y utilizacion la potenciación, la radicación y la logaritmación para representar situaciones matemáticas y no matemáticas y para resolver problemas.

1. CONCEPTOS BÁSICOS DE LAS MATEMÁTICAS

Para tener una mejor comprensión de la estadística y su metodología es muy importante tener claridad de los conceptos básicos de las matemáticas que ella involucra. Pues, ayudan en gran medida en el desarrollo de las operaciones necesarias en la aplicación de los conceptos y teoremas estadísticos.

Es importante tener en cuenta que las matemáticas proporcionan las técnicas, las herramientas y la tecnología que permiten hacer estadística y analizar la información proveniente de un conjunto de datos o de una muestra representativa tomada de una población en particular.

Por lo expuesto anteriormente, en este capítulo se presenta un pequeño tratado de matemáticas, un cuerpo de conceptos fundamentales y de métodos básicos matemáticos para formar una visión general de su importancia en la estadística y en el desarrollo de la misma.

1.1 CONJUNTO

En matemática se entiende por conjunto los elementos que poseen unas características o propiedades propias , también se expresa como "*una asociación mental de todo en ciertos objetos diferentes que pueden ser abstractos o reales*"[1] *(Punto de Vista Teoría Intuitiva)*; Ejemplos de conjuntos son: Los días de la semana, los meses del año, las letras del alfabeto, entre otros.

[1] Definición de G. Cantor (1845-1918), Creador de la Teoría de Conjuntos

Los elementos de un conjunto se designaran con letra minúscula como •, •, •, y los conjuntos se designan con letras mayúsculas como •, •, •; si el elemento • pertenece al conjunto •, su notación matemática es la siguiente • • •. Si el elemento • no pertenece al conjunto • se denota, • • •.
Si los elementos de un conjunto • son también elementos de un conjunto •, se dice que • es ***subconjunto*** de •, y su notación es • • •. Si se verifica que • • • y • • •, entonces • es ***subconjunto propio*** y se escribe • • •

1.1.1 Tipos de Conjuntos

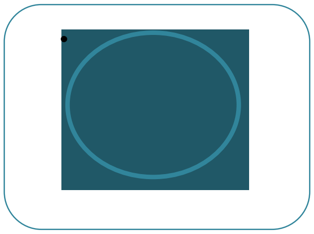
Si se posee un número de elementos finitos en conjunto se conoce como **conjunto finito**, si se desconoce el número de elementos de un conjunto, en otras palabras que sus elementos sean ilimitados este tipo de conjuntos es llamado **conjunto infinito.**

El **conjunto acotado** es aquel que define una relación de orden la cual posee una cota superior e inferior; cuando se establece una relación de orden entre los elementos de un conjunto este es conocido como **conjunto ordenado**; el conjunto que solo posee un sólo elemento se llama **conjunto unitario**, el conjunto que contiene todos los elementos se le conoce como **conjunto universal** (**U**), el conjunto que no tiene elemento alguno es llamado **conjunto vacio** y se representa por el símbolo **Ø**.

1.1.2 Representación de conjuntos con Diagramas de Venn.

Este diagrama es usado con frecuencia para mostrar las relaciones entre conjuntos.

Figura 1.1 Diagrama de un conjunto

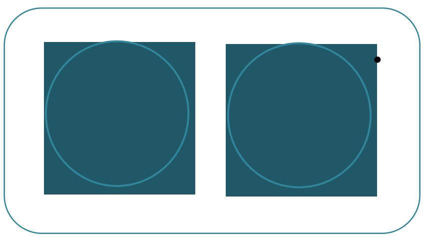


El diagrama de Venn de la figura 1.1, ilustra un conjunto •

.

Fuente: Los Autores

Figura 1.2 Diagrama de conjuntos disyuntos



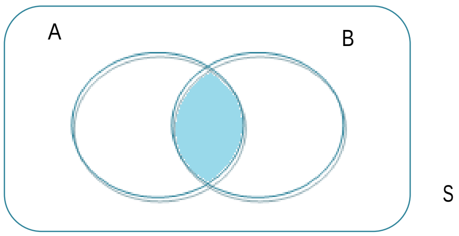
El diagrama de Venn de la figura 1.2, ilustra dos conjuntos • y • que no tienen elementos comunes.

Este tipo de conjuntos se le conoce como conjuntos disyuntos.

Fuente: Los Autores

Matemáticamente, se simboliza: • • • = •. Es decir que el conjunto A y el conjunto B no tienen ningún elemento común. En estadística dos o más conjuntos disyunto representan conjuntos mutuamente excluyentes.

Figura 1.3 Diagrama de Intersección de conjuntos



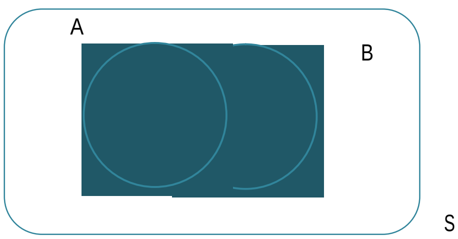
El diagrama de Venn de la figura 1.3, ilustra dos conjuntos y que poseen al menos un elemento en común.

A esta relación de conjuntos se le conoce como intersección de conjuntos.

Fuente: Los Autores

Matemáticamente intersección de dos conjuntos se representa: . Es decir, para todo .

Figura 1.4 Diagrama de unión entre conjuntos



El diagrama de Venn de la figura 1.4, ilustra dos conjuntos y con características similares.

A este tipo de conjunto, se le conoce como unión de conjuntos.

Fuente: Las Autoras

La unión de dos conjuntos corresponde a la unificación de elementos de dos o más conjuntos que puede conformar un nuevo conjunto, en el cual los elementos del nuevo conjunto corresponden a los elementos de los conjuntos originales.

Matemáticamente, si es decir para todo x,

Otro método para representar conjuntos es a través de llaves, Así:

1.1.3 Conjunto de los números reales

Para iniciar la construcción de los números reales se comienza con el conjunto de los ***números naturales*** también llamados números de conteo y del que hacen parte de los números enteros positivos. Estos se representan:

$$• = \{1, 2, 3, 4, \ldots , • \}$$

Cuando al conjunto de los números naturales se le agrega el cero, obtenemos los números ***enteros positivos:***

$$•^{•} = \{0, 1, 2, 3, 4, \ldots , • \}$$

Los números enteros comprenden dos conjuntos**:** los números enteros positivos y los ***números enteros negativos:***

$$•^{•} = \{• • , \ldots , • 3, • 2, • 1\}$$

Y la notación de los números enteros simbólicamente es la siguiente:

$$• = \{• • , \ldots , • 3, • 2, • 1, 0, 1, 2, 3, 4, \ldots , • \}$$

El conjunto de los **números racionales** son aquellos de la forma $\frac{•}{•}$, donde • y • son números enteros, con • • 0. La notación de conjunto se escribe**:**

$$• = \{• / • | • • • • • • • • • • • • •, • • 0\}$$

No solo estos números forman el conjunto de los números, también le pertenecen el conjunto de los números irracionales (• •). Así, el conjunto de los números reales (•), se define como la unión de los racionales con los irracionales

$$• = • • ••$$

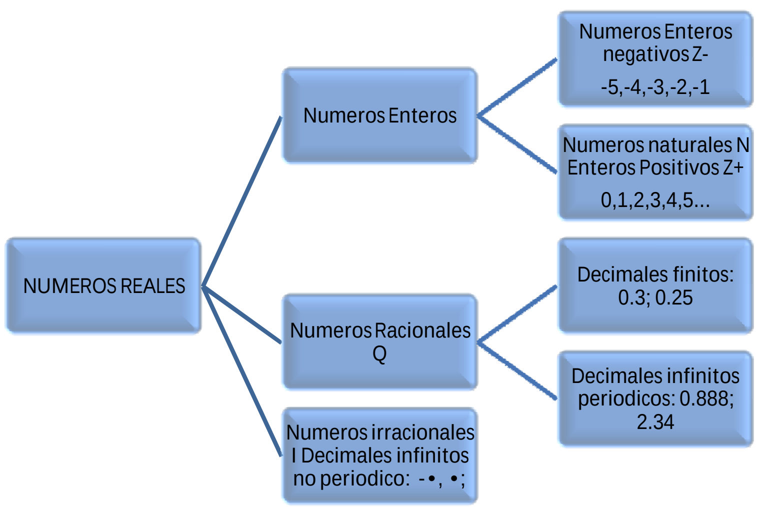
Ejemplo de números **irracionales** ( ), son – – .

Nótese que está contenido en , pues cada entero se puede describir de la forma – ; con .

Ejemplo 1:

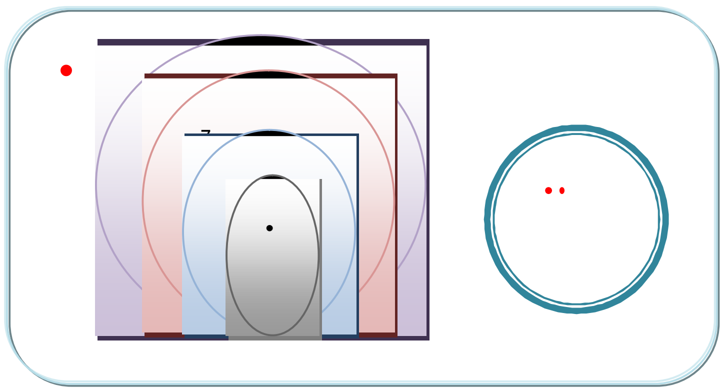
El numero se puede escribir en la forma – – , aunque los números enteros estén contenidos en los números racionales, los números racionales no están contenido en los números enteros, pues las fracciones como – — no son de tipo entero.

Diagrama 1.1 Conjunto de los números reales.



Fuente: Los Autores

Figura 1.6. Conjunto de los números reales.



Fuente: Los Autores

En el grafico anterior se muestra la relación existente entre los diferentes conjuntos numéricos.

1.1.4 El Valor Absoluto de un Número Real

El valor absoluto ó también llamado *módulo* de un número real se encuentra definido de la siguiente forma:

- Siendo • un número real positivo, su valor absoluto es el mismo número.
- Si • es negativo, su valor absoluto es el inverso aditivo, •• (que es positivo).
- Si • es 0, su valor absoluto es 0[2].

[2] Kearny patricia A. Fundamento de matemáticas para estadística elemental: Texto programado. Editorial Limusa. México. 1973.

En resumen:

|•| = • •, •• • • 0 •
•• , •• • < 0

Cabe resaltar que el valor absoluto de un número real ‘••siempre va hacer de tipo mayor ó igual a cero, pero **nunca** negativo.

Ejemplo 2:

|9|= 9
|-9|=9
|0|=0

Propiedades del valor absoluto.

|•| • 0	No negatividad
|•| = 0 • • = 0	Definición positiva
|••| = |•||•|	Propiedad Multiplicativa
|• + •| • |•| + |•|	Propiedad Aditiva
|• •| = |•|	Simetría
|• • •| = 0 • • = •	Identidad de indiscernibles
|• • •| • |• • •| + |• • •|	Desigualdad Triangular
•÷• = |•|/|•| •• • • 0	Preservación de la división

El valor absoluto es útil para representar resultados, como: la distancia recorrida por una persona u objeto, las piezas o productos fabricados por una maquina en un periodo de tiempo, la deuda financiera de una persona, la cantidad de unidades faltantes de determinado producto para satisfacer la demanda, ya que son variables que nunca podrán tomar valores negativos.

1.1.5 Representación de los números reales como decimales

Cualquier número real puede escribirse como un decimal finito o infinito; por ejemplo~~,~~ teniendo un número real de forma fraccionaria $4/3$, se puede representar mediante el decimal: 1,33333.

En muchos textos se puede encontrar el decimal de la forma: $1,\overline{3}$ esta barra sobre el 3 significa que el mismo se repite indefinidamente. Otro ejemplo es el numero ½ se representa mediante el decimal 0,5. En este caso 5 se repite solo una vez, por tal razón el número se considera como un decimal finito.

Tabla 1.1. Representación decimal de los conjuntos.

Conjunto	Descripción	Ejemplos	Representación decimal
Números naturales	Números positivos	1, 2, 3,…	**Decimales Finitos**
Números enteros no negativos	Números positivos y conjunto vacio	0,1,2,3,…	**Decimales Finitos**
Números enteros	Números positivos, números negativos y conjunto vacio	..-3,-2,-1,0,1,2,3,..	**Decimales Finitos**
Números Racionales	Números que se pueden escribir como a/b, donde a y b son enteros y b • 0	-3, -3/4, -0.222, 0.5/6,	**Decimales Finitos o infinitos periódicos**
Números irracionales	Números que no se pueden escribir como a/b, donde a y b son enteros y b • 0	$\sqrt{2}$, $\sqrt{3}$, • 1.414213... 1.732050...	**Decimales infinitos no periódicos**
Números reales	Números racionales e irracionales	Todos los anteriores	**Todo tipo de decimales**

Fuente: Kearny patricia A. Fundamento de matemáticas para estadística elemental: Texto programado. 1973.

Un número irracional se representa como decimal, este no es periódico ni finito, por ejemplo:

$$\sqrt{2} = 1,41421\ldots \quad \bullet\ \bullet = 3.14159\ldots$$

1.1.5 Operaciones con números reales

Las operaciones usuales con números reales son:

- La adición ó suma, notada con el símbolo más (+), permite operar dos números cualesquiera • • • para conseguir su suma, que se denota • + •.
- La multiplicación ó producto, notada con el símbolo (•), permite operar dos números cualesquiera *a* y *b* para formar su producto, • • • (*ó* • •).
- La sustracción o resta se define en términos de la suma. Así:

$$\bullet \; \bullet \; \bullet = \bullet + (\bullet \, \bullet)$$

Donde – • es el inverso aditivo de •.

- De manera análoga la operación de división se define en términos de la multiplicación. Debido a que el inverso multiplicativo de un número real •, el cual es distinto de cero; es $\frac{\bullet}{\bullet}$, que también se escribe $\bullet^{\bullet\,\bullet}$. Entonces,

$$\bullet \; \bullet \frac{1}{\bullet} \bullet$$

Se escribe $\frac{\bullet}{\bullet}$, y se dice que a se divide entre •. Ejemplo:

$$6 \bullet \frac{1}{3} \bullet = \frac{6}{3}$$

Hay que tener en cuenta que cero no tiene inverso multiplicativo, pues la división entre cero no está definida.

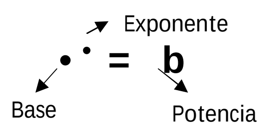
Tabla 1.2. Axiomas de los números reales

Reglas		Ejemplo
Adición		
1. • + • = • + •	Ley conmutativa de la suma	*5 + 2 = 2 + 5*
2. • + (• + •) = (• + •) + •	Ley asociativa de la suma	*(3+5) + 2 = 3 + (5+2)*
3. • + • = •	Ley del neutro de la suma	*2+0=2*
4. • + (• •) = •	Ley del inverso de la suma	*7 + (-7)=0*
Multiplicación		
1. • • = • •	Ley conmutativa de la multiplicación	*6 • 2 = 2*• 6
2. •(• •) = (• •)•	Ley asociativa de la multiplicación	6(2 • 4) = (6 • 2)4
3. • • • = • • •	Ley del neutro de la multiplicación	6 • 1 = 1 • 6
4. • • $\frac{•}{•}$ = • (• • •)	Ley del inverso de la multiplicación	3 • $\frac{•}{•}$ = 1 (3 • •)
Distributivas		
1. •(• + •) = •• + ••	Ley distributiva de la multiplicación respecto de la suma	3(4 + 5) = 3 • 4 + 3 • 5

Fuente: Kearny patricia A. Fundamento de matemáticas para estadística elemental: Texto programado. 1973

1.1.6. Potenciación de números enteros

La potenciación es una operación que consiste en multiplicar un número, llamado base, por si mismo varias veces según lo indique el exponente, el resultado de la potenciación se llama potencia.



Cuando el exponente es un numero entero positivo, la potenciación es multiplicar el número por si mismo las veces que lo indique el exponente

$$\underbrace{a * a * a * ... * a}_{n\ veces} = b$$

$$\underbrace{3*3*3*3}_{4\ veces} = 81$$

En el caso que el exponente es un número entero negativo, la potencia es la fracción inversa de la base con exponente positivo

$$\bullet^{\bullet\,\bullet} = \bullet\frac{\bullet}{\bullet}\bullet^{\bullet}$$

$$\underbrace{\frac{\bullet}{\bullet} * \frac{\bullet}{\bullet} * \frac{\bullet}{\bullet} * \frac{\bullet}{\bullet}}_{4\ \text{veces}} = \frac{\bullet}{\bullet\bullet}$$

1.2.6.1 Propiedades de la Potenciación

- *Producto de una potencia de igual base*

El resultado es otra potencia de la misma base y de exponente igual a la suma de los exponentes de los factores, es decir:

$$2^{\bullet} \bullet 2^{\bullet} \bullet 2^{\bullet} = 2^{\bullet\bullet\bullet\bullet\bullet} = 2^{\bullet\bullet}$$

Simbólicamente: $\bullet^{\bullet} \bullet \bullet^{\bullet} \bullet \bullet^{\bullet} \bullet \ldots \bullet^{\bullet} = \bullet^{\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet}$ Si $\bullet \in$ **R**, x, y, z,…, *n* $\in$ **N**

- *Cociente de potencias de igual base.*

Cuando se tiene un cociente de igual base elevadas a diferentes exponentes, el resultado será otra potencia de la misma base y de exponentes igual a la diferencia entre el exponente del numerador y el exponente del denominador, por lo tanto:

$$\frac{7^{\bullet}}{7^{\bullet}} = 7^{\bullet\bullet\bullet} = 7^{\bullet}$$

Simbólicamente: $\frac{\bullet^{\bullet}}{\bullet^{\bullet}} = \bullet^{\bullet\bullet\bullet}; \bullet\bullet\bullet\ \bullet\ \bullet\ \bullet$, m, n $\bullet\ \bullet; \bullet\ \bullet\ 0, \bullet > \bullet$

Es importante resaltar que esto se cumple siempre y cuando el exponente del numerador es mayor o igual al exponente del denominador.

- *Potencia de una potencia.*

Es una potencia de la misma base, y de exponente igual al producto de los exponentes que haya en la expresión, luego: (5•)• = 5• × • = 5••

Simbólicamente: (• •)• = • • × • = • • •,••• • • •, •,• • •.

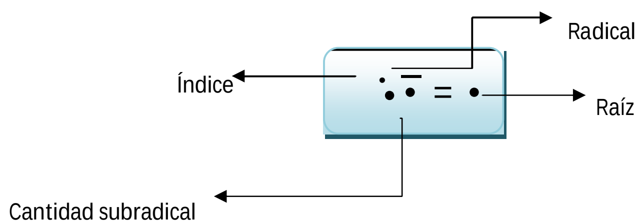
- *Potencia de un producto*

Es igual al producto de dichas potencias, es decir: (5 × 2)• = 5• × 2• = 5•2•

Simbólicamente: (• × •)• = •• × •• = ••••,••• •,• • •,• • •.

1.2.7 La radicación de números enteros

La radicación es la operación inversa a la potenciación ó en otras palabras cuando se necesita conocer el valor de un número • expresado de la forma •• .



- Cuando hablamos de raíz cuadrada no se escribe índice, y el número • que se encuentra debajo del símbolo radical es llamado *radicando* o *cantidad subradical* y b es la raíz.

- Las raíces que presentan índice par tienen dos posibles soluciones pero se considera solo el positivo por la regla de signos
- Cabe resaltar que las raíces de base negativa e índice par no tienen solución, ya que ningún número entero elevado a un exponente par da como resultado un número negativo

Ejemplo 3:

En un estudio realizado a 20 hombres acerca de las edades en las cuales éstos inician su vida laboral se obtiene los siguientes resultados:

Tabla 1.3. Datos de entre hombre y su edad en la cual inician a laborar

Hombre encuestado	***Edad en la cual inicia a trabajar***	***Hombre encuestado***	***Edad en la cual inicia a trabajar***
1	*16*	***11***	*20*
2	*17*	***12***	*16*
3	*15*	***13***	*24*
4	*18*	***14***	*15*
5	*17*	***15***	*14*
6	*20*	***16***	*12*
7	*15*	***17***	*21*
8	*17*	***18***	*18*
9	*18*	***19***	*19*
10	*19*	***20***	*17*

Fuente: Los Autores

Si se desea saber la edad aproximada en la cual los hombres inician su vida laboral, se debe sumar todas las edades y dividir entre el número de hombres esto es:

$$• = \frac{348}{20} = 17{,}4 \text{ •ñ••}$$

Este resultado de nota el promedio

Con formato: Normal

1.2.8 El signo igual

El *signo igual* (=) entre dos expresiones asegura que las dos expresiones son nombres o descripciones exactas del mismo objeto. El signo (•) significa “diferente a” o “no es igual a”, y sirve para expresar la diferencia de los objetos.

Ejemplo 4:

X=3
y• 5
Número de estudiantes del grupo G = Número de estudiantes del grupo K

1.2.9 Desigualdad

Una desigualdad es una expresión matemática en la que ambos miembros de una afirmación no son necesariamente iguales, por ejemplo se puede decir 56 es mayor que 15, o que 4 es menor que 16. Para denotar esta situación se usan los símbolos > que representa “…es mayor que…”, < para representar “…es menor que…”, • que significa “es menor que o igual a”, de manera similar, • significa “es mayor que o igual a”. En resumen:

• < •	• es menor que •
• > •	• es mayor que •
• • •	• es menor que o igual a •
• • •	• es mayor que o igual a •

Con formato: Punto de tabulación: 6,25 cm, Izquierda

Ejemplo 5:

El Instituto de medicina legal y ciencias forenses en acompañamiento con el sistema de información de desaparecidos y cadáveres y el sistema de información para el análisis de la violencia y accidentalidad en Colombia realiza un estudio para determinar los muertos y lesionados en accidente de tránsito según sexo e hipótesis causal en Colombia en el primer semestre de 2008, la información de dicho estudio refleja que:

Tabla 1.4. Muertos y lesionados en accidente de tránsito según sexo e hipótesis causal en Colombia en el primer semestre de 2008

Circunstancia	Mujeres	Hombres
Violación normas de otras normas de transito	56	198
Otros	112	363
Exceso de velocidad	78	324
Violación de normas de transito peatones	27	89
Embriaguez (alcohólica y no alcohólica)	13	174
Contravía	9	27
Posibles fallas mecánicas	33	93
Irrespeto de los semáforos	4	9
Malas condiciones en la vía	18	46
Malas condiciones ambientales	7	14
Subtotal	357	1337
Sin información	712	2884

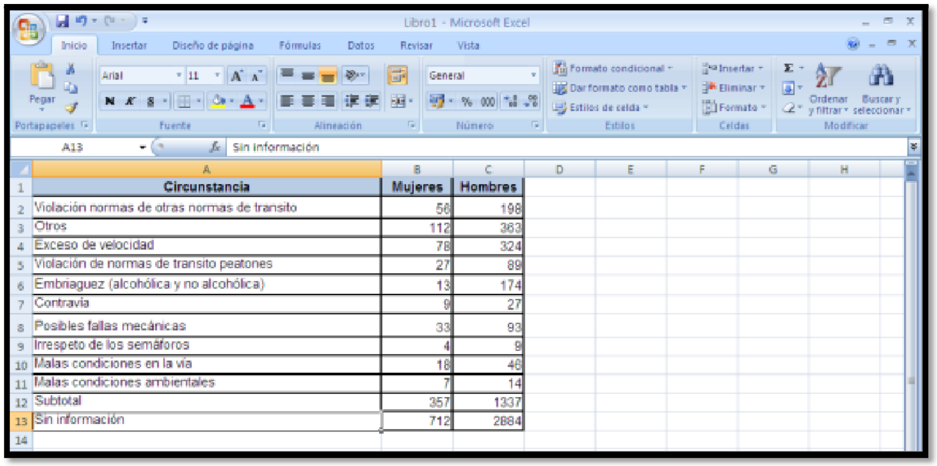
Fuente: Tomado de http://www.medicinalegal.gov.co- Muertes y lesiones por accidente de transito

De acuerdo a la tabla anterior es posible afirmar que: El número de accidentes causados por posibles fallas mecánicas en el vehículo (>) "*es mayor que*" el número de accidentes causados por malas condiciones en la vía.

Para realizar el ejercicio a través de Excel se deben realizar los siguientes pasos:

1. Se introducen los datos del ejemplo al libro Excel, como se muestra en la figura siguiente:

Figura 1.7 Datos en Excel

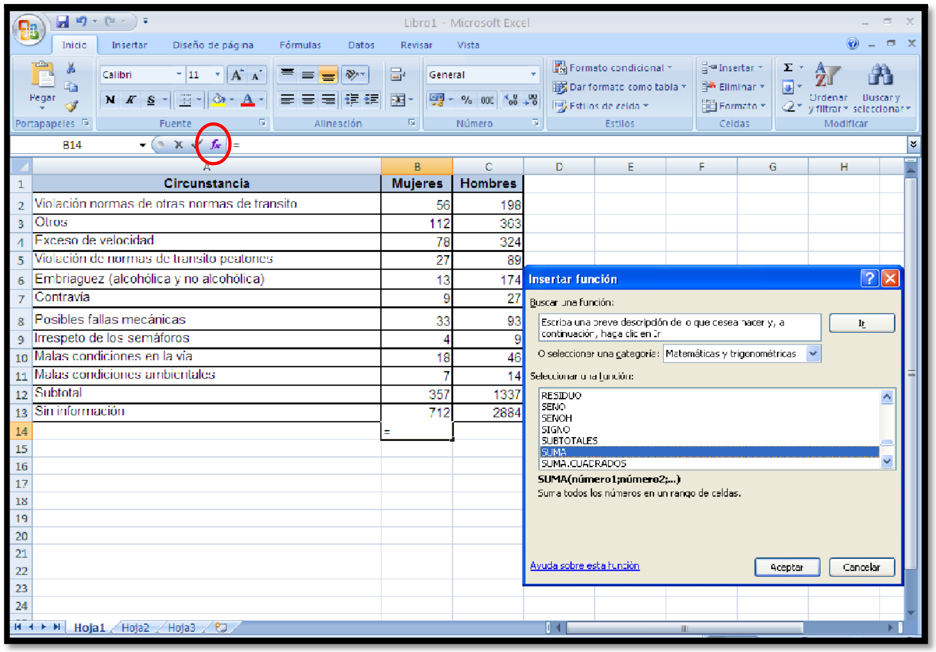


Circunstancia	Mujeres	Hombres
Violación normas de otras normas de transito	56	198
Otros	112	363
Exceso de velocidad	78	324
Violación de normas de transito peatones	27	89
Embriaguez (alcohólica y no alcohólica)	13	174
Contravía	9	27
Posibles fallas mecánicas	33	93
Irrespeto de los semáforos	4	9
Malas condiciones en la vía	18	46
Malas condiciones ambientales	7	14
Subtotal	357	1337
Sin información	712	2884

Fuente: Los Autores

2. Para obtener la sumatoria de las edades, se ingresa a la función (*$f_{x.}$)* Se despliega un cuadro de dialogo llamado *"Insertar función"* y se selecciona una categoría, para llevar a cabo el ejemplo se debe escoger la función Matemáticas y Trigonométricas e inmediatamente aparecen diferentes funciones, de las cuales se tomara la función **suma** y se da clic en aceptar o se presiona la tecla *"Enter"*.

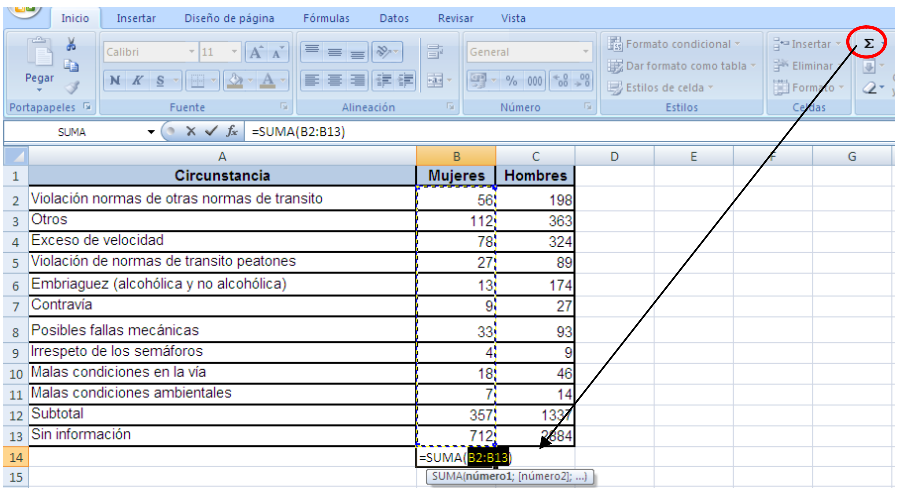
Figura 1.8 Excel Aplicado al ejercicio



Fuente: Los Autores

3. En el cuadro argumentos de la función, se da clic en el icono señalado en la siguiente figura:

Figura 1.9 Excel Aplicado al ejercicio



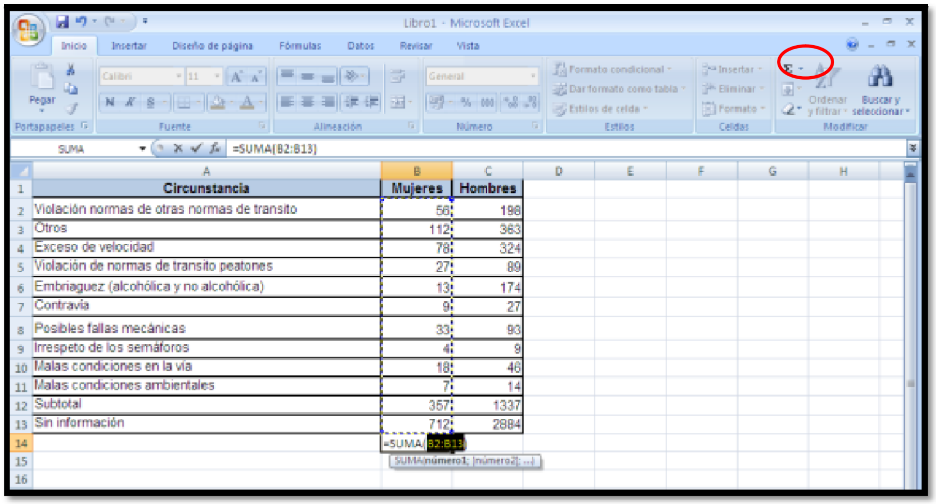
	A	B	C
1	Circunstancia	Mujeres	Hombres
2	Violación normas de otras normas de transito	56	198
3	Otros	112	363
4	Exceso de velocidad	78	324
5	Violación de normas de transito peatones	27	89
6	Embriaguez (alcohólica y no alcohólica)	13	174
7	Contravía	9	27
8	Posibles fallas mecánicas	33	93
9	Irrespeto de los semáforos	4	9
10	Malas condiciones en la vía	18	46
11	Malas condiciones ambientales	7	14
12	Subtotal	357	1337
13	Sin información	712	2884
14		=SUMA(B2:B13)	

Fuente: Los Autores

Y se introduce el rango de valores de la columna “mujeres”, que corresponden al número de mujeres que se accidentaron por diferentes circunstancias y se da clic en aceptar, otra manera de escoger el rango es escribiendo la dirección de la celda de inicio y la celda donde se encuentra el ultimo datos. Por ejemplo: *(B3:B13)*. De igual manera se puede hacer para obtener la sumatoria del número de hombres accidentados.

Otra forma de realizar la operación de la sumatoria más rápida es a través del icono que aparece en la parte superior de la barra de herramientas indicado con ó sumatoria, para esto solo hay que dar clic sobre el icono y seleccionar el rango de valores que se desea sumar.

Figura 1.10 Excel Aplicado al ejercicio



Circunstancia	Mujeres	Hombres
Violación normas de otras normas de transito	56	198
Otros	112	363
Exceso de velocidad	78	324
Violación de normas de transito peatones	27	89
Embriaguez (alcohólica y no alcohólica)	13	174
Contravía	9	27
Posibles fallas mecánicas	33	93
Irrespeto de los semáforos	4	9
Malas condiciones en la vía	18	46
Malas condiciones ambientales	7	14
Subtotal	357	1337
Sin información	712	2884
	=SUMA(B2:B13)	

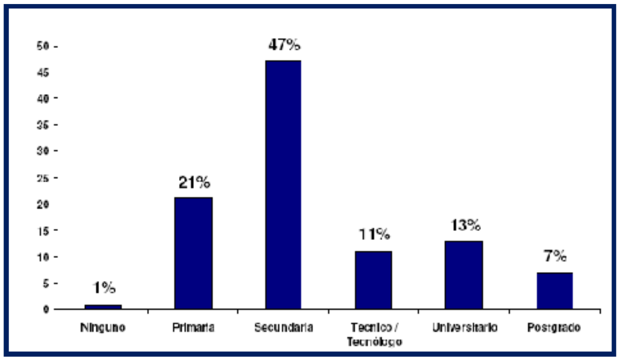
Fuente: Los Autores.

De manera general, se pueden realizar diversas operaciones que se encuentran en el cuadro de dialogo de insertar función, en esta se encuentran diferentes categorías tales como matemáticas y trigonométricas, estadísticas, financieras, ingeniería, lógicas, entre otras; y cada una dispone de sus respectivas funciones.

Ejemplo 6:

El proyecto Cartagena Como Vamos en el año de 2007 realizó una encuesta para evaluar las expectativas de estudiantes de último año de educación media en Cartagena, tomando como muestra a los estudiantes de colegios públicos y privado de la cuidad. En su boletín de informe de resultados de la encuesta emitida en julio de 2008, se encuentra una pregunta que relaciona el nivel educativo de los padres, el resultado se describe a continuación:

Figura 1.11. Nivel educativo de los padres



Fuente: Proyecto Cartagena como vamos, 2007

A partir de la grafica anterior es importante destacar que el porcentaje de padres que cursan hasta el nivel básica primaria es menor que el porcentaje de padres que cursan hasta la secundaria.

1.3 SUMATORIA Y PROPIEDADES DE LAS SUMATORIAS

El operador matemático que nos permite representar sumas muy grandes, ya sea de n o incluso infinitos sumandos está expresado con la letra griega sigma (•).

La sumatoria se representada por un símbolo

Donde Σ es una letra griega correspondiente a S, llamada *Sigma*.

Donde:

i Elemento genérico de la sumatoria
n : Limite superior de la sumatoria
Σ : Letra sigma e indica la sumatoria
i=1: índice o límite inferior de la suma

Teniendo en cuenta que la suma es una operación de los números enteros, en este sentido, la sumatoria es la suma de datos de una variable (*x*): $x_1, x_2, x_3, \ldots, x_n$.

Ejemplo 7:

Del ejemplo 1.1 se quiere obtener el total de:

- Mujeres accidentadas por circunstancias conocidas
- Mujeres accidentadas en el primer semestre de 2008
- Hombres accidentados por circunstancias conocidas
- Hombres accidentados en el primer semestre de 2008

En este caso, tomamos x_i, donde i le corresponden los valores del número de accidentes, reemplazando cada x_i por su valor correspondiente, la solución de dicha sumatoria sería:

- Mujeres accidentadas por circunstancias conocidas

  ••
  • •. = 56 + 112 + 78 + 27 + 13 + 9 + 33 + 4 + 18 + 7 = 357
  •••

- Mujeres accidentadas en el primer semestre de 2008, esto es:

  ••
  • •. + •• • ••• •• • •••••• • sin ó•
  •••

  Total de mujeres accidentadas = 357+712 =1426

- Hombres accidentados por circunstancias conocidas

  ••
  • •. = 198 + 363 + 324 + 89 + 174 + 27 + 93 + 9 + 46 + 14 = 1337
  •••

- Hombres accidentados en el primer semestre de 2008, también se puede expresar como:

•• • •. + •• • •••• •• •• • •••• sin ó•
•••

Total de Hombres accidentados = 1337 + 2884 = 5558

1.2.1 Propiedades de la sumatoria.

- La sumatoria de una constante de 1 a *n* es igual a *n* veces la constante.

• • = • + • + • + • = ••

- La sumatoria del producto de una constante por una variable es igual a la constante por la sumatoria de la variable.

• ••. = ••. + ••. + • + ••. = •(•. + •. + • + •.)

• ••. = • • •.

- La sumatoria de los valores de una variable mas una constante es igual a la sumatoria de la variable mas *n* veces la constante.

• (•. + •) = (•. + •) + (•. + •) + • + (•. + •) = •. + •. + • + •. + • + • + • + •

•(•. + •) = • •. + ••

Corolario

• (•. • •) = • •. • ••

- Las sumatoria de una constante, cuyo límite inferior es diferente de 1 se expresa:

•
• • = (• • • + 1).
•••••
•, ••••• • =

1.2.2. Fórmulas empleadas en la sumatoria

En algunos casos se puede emplear la letra • en vez de x_i, de forma general:

•
• •
•••

Donde, • es el elemento genérico de la sumatoria.

Ejemplo 8:

Un grupo de estudiantes de Psicología de la UTB desea comparar los resultados de la encuesta del ejemplo 1.4, para confirmar el promedio de edad en la cual los hombres inician su vida laboral. Para ello escoge 20 muestras de la población (hombres de la ciudad de Cartagena), pero deciden escoger el mismo número de muestra para cada población, los estudiantes quieren saber ¿Cuantos hombres encuestarán?

Para obtener el número de hombres a encuestar es necesario aplicar la formula anterior. Así:

•::.• = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 14 + 15 + 16 + 17 + 18 + 19 + 20 = 210

En conclusión, se encuestarán a 210 hombres.

Otras operaciones de sumatoria simple son:

$$\bullet \dot{\vdots}_{\cdots} \bullet^{\bullet} ; \quad \bullet \dot{\vdots}_{\cdots} \bullet^{\bullet} ; \quad \bullet \dot{\vdots}_{\cdots} \bullet^{\bullet}$$

Estas son las formulas sencillas que proporcionan el valor de la suma de • números comprendidos entre 1 y •, inclusive[3]:

$$\vdots_{\cdots} \bullet = \frac{\bullet(\bullet + 1)}{2}$$

$$\vdots_{\cdots} \bullet^{\bullet} = \bullet\frac{\bullet(\bullet + 1)}{2}\bullet^{\bullet}$$

Ejemplo 9:

$\bullet \dot{\vdots}_{\cdots} \bullet$ = 28; aplicando la formula: $\bullet \dot{\vdots}_{\cdots} \bullet = \frac{\bullet(\bullet\bullet\bullet)}{\bullet}$

$$\vdots_{\cdots} \bullet = \frac{7(7 + 1)}{2} = 28$$

Para la fórmula: $\bullet \dot{\vdots}_{\cdots} \bullet^{\bullet} = \bullet\frac{\bullet(\bullet\bullet\bullet)}{\bullet}\bullet^{\bullet}$

La sumatoria de las 7 encuestas elevadas al cubo.

$$\vdots_{\cdots} \bullet^{\bullet} = \bullet\frac{7(7 + 1)}{2}\bullet^{\bullet} = 784$$

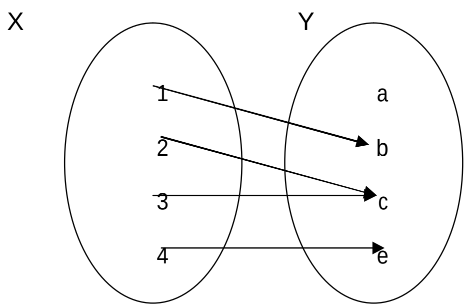
[3] Adaptado de Ciro Martínez

1.3 FUNCIÓN MATEMÁTICA

Una función matemática es una aplicación entre dos conjuntos numéricos de forma que a cada elemento del primer conjunto le corresponde uno y sólo un elemento del segundo conjunto:

f : X Y

x -> y = f(x)



La Función de X en Y: según la condición de existencia asegura que de cada elemento sale alguna flecha y la de unicidad que sólo sale una.

Los elementos de X se llaman origen y los elementos de Y se llaman imagen,

Se debe cumplir que X se encuentre relacionado con los elementos de Y considera una relación de conjuntos la correspondencia de un elemento con otro, o cuando existe una conexión entre ellos. Es decir, la conexión de un elemento de un conjunto con elementos de otro conjunto.

Una función se denota por la letra *f* y se puede representar:

• • • → • *que se lee " función de A en B"*

Con formato: Interlineado: 1,5 líneas

Entonces, se puede definir una función matemática como una relación entre elementos de conjuntos. Donde los elementos se les conocen con el nombre de variables, debido que al describir la relación entre ellos, pueden variar.

Con formato: Sangría: Izquierda: 0 cm, Primera línea: 0 cm

Una función matemática no solo depende de la asociación de los conjuntos • y •, sino también de la definición de los mismos en la función. Ya que si se cambian los elementos de • o •, la función cambia.

Ejemplos de funciones pueden ser:

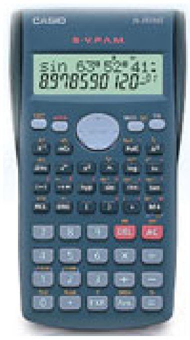
- El costo de un pan entre diferentes panes
- El costo de una llamada telefónica que depende de la duración de la misma
- El salario de un trabajador que depende del total de horas trabajadas.

1.4 USO DE LA CALCULADORA

La calculadora es una herramienta que sirve para resolver problemas que requieran cálculos grandes, o también puede ser de utilidad para resolver algunos problemas de ingenio. La calculadora consta principalmente de un teclado, con el cual introduces los números --y operaciones que deseas realizar, y de una pantalla o "display", en la cual aparecen los números que teclas y los resultados de las operaciones que realiza la calculadora.

Teclas marcadas con los números dígitos:

Con formato: Interlineado: 1,5 líneas, Punto de tabulación: 3 cm, Izquierda



Teclas marcadas con los números dígitos:

9 8 7 6 5

4 3 2 1 0

Teclas marcadas con funciones especiales:

Tecla para ejecutar las operaciones: =

Tecla para colocar el punto decimal: .

Teclas para ejecutar las operaciones básicas:

Suma: +

Resta: -

Multiplicación: x o *

División: ÷ o /

Para realizar algunas operaciones como potencias, raíces y recíprocos, las calculadoras disponen de las siguientes teclas:

Para obtener el cuadrado de un número $\boxed{x^2}$

Para obtener la raíz cuadrada de un número

Para obtener el resultado de un número elevado a cualquier exponente

Para obtener la raíz n-sima de un número

Las anteriores teclas es posible acceder por medio de la tecla función $\boxed{\text{shift}}$

Ejemplo de cómo se usan estas teclas:

De acuerdo a una encuesta realizada en la que se evalúa las calificaciones de un grupo de estudiantes de estadística, se tiene que la varianza de las calificaciones es 2, se desea conocer el patrón de variación de dicha variable, esto es la varianza. Conceptualmente:

La desviación estándar=

Por lo cual, la desviación estándar es igual a $\sqrt{4}$ que es lo mismo que 2. Para hacer este cálculo, inicialmente se oprime la tecla que denota la operación raíz cuadrara y se introduce el numero. Sin embargo, otros tipos de calculadora esta operación se realiza de manera inversa, primero se introduce el número y luego la tecla que muestra la raíz cuadrada.

En caso tal, que el dato que se proporcionara en el estudio fuera la desviación estándar, en vez de la varianza, se puede hallar la varianza, elevando esta el cuadrado. De este modo:

La varianza= (La desviación estándar)^2

El procedimiento en la calculadora para hacer la operación de elevar el cuadrado de cualquier número se realiza de manera similar a la operación de la raíz cuadrada. Se introduce el dato y se oprime la tecla que muestra el cuadrado de un número y se obtiene el resultado deseado.

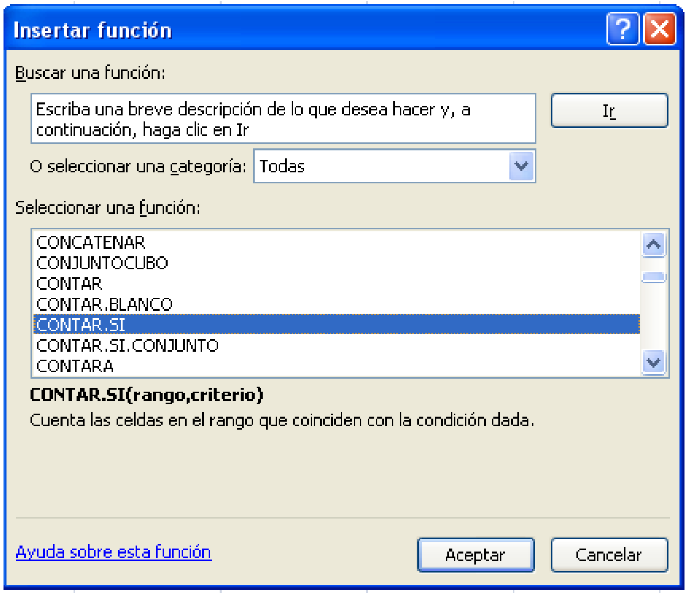
Otras de las operaciones que se puede realizar en la calculadora es elevar el valor de una variable a una potencia "x" específica. Por ejemplo si se desea obtener el resultado de 4^5, primero se escribe el número 5, se oprime la tecla $\boxed{\wedge}$ o en otras calculadoras se encuentra de la forma

Esta tecla sirve para obtener la raíz específica x de un número (y). Por ejemplo, Para hallar la raíz quinta de 256 en la calculadora, inicialmente se introduce el índice en este caso es 5, luego se oprime la tecla para así introducir la cantidad llamada radical $\overline{256}$ obteniendo como resultado 3,031.

APLICACIÓN EXCEL

1.5 FUNCIONES BÁSICAS DE EXCEL

Figura 1.12 Aplicaciones en Excel



Fuente: Los Autores.

Funciones Estadísticas: Excel ofrece diversas funciones estadísticas, tales como:

- CONTAR.SI: Por medio de esta función se puede contar la cantidad de veces que se repite un carácter o una palabra que concuerde con las características descritas.
- DESVESTA: Por medio de esta función se puede calcular el valor de la desviación estándar de una muestra.
- FRECUENCIA: calcula el número de veces que se repite un valor en la muestra obtenida
- MAX: a partir de esta función se obtiene el valor máximo dentro de la muestra.

- MIN: A partir de esta función se obtiene el valor mínimo dentro de la muestra.
- PROMEDIO: Corresponde a la media aritmética de una muestra, el promedio se obtiene de datos numéricos.
- VAR: Esta función sirve para hallar la varianza en una muestra.

Funciones matemática y trigonométrica:

- ALEATORIO: A partir de esta función se puede calcular un número aleatorio entre 0 y 1.
- FACT: Por medio de esta función se puede calcular el factorial de un número.
- EXP: Calcula el valor de la siguiente expresión matemática ••; donde • es la representación de una constante, generalmente conocida como Euler, y la letra *n* es el valor a ingresar
- LN: A través de esta función se obtiene el logaritmo natural de un número cualquiera
- LOG: Sirve para hallar el logaritmo de un número, definiendo la base.
- POTENCIA: Calcula el resultado de un número elevado a una potencia cualquiera.
- PRODUCTO: Realiza la multiplicación de dos o más números
- RAÍZ: Calcula la raíz cuadrada de un número cualquiera.
- REDONDEAR: Aproxima un numero a la cantidad de decimales que se desea.
- REDONDEAR.MAS: Aproxima un numero a la cifra decimal siguiente.
- REDONDEAR. MENOS: Aproxima un numero a la cifra decimal anterior.
- SUMA: Realiza la operación de suma en los rangos establecidos.
- TRUNCAR: Transforma los decimales en numero enteros, o fracciona la parte decimal del numero

Con formato: Fuente: Negrita, Español (Colombia)

APLICACIÓN STATGRAPHICS CENTURION

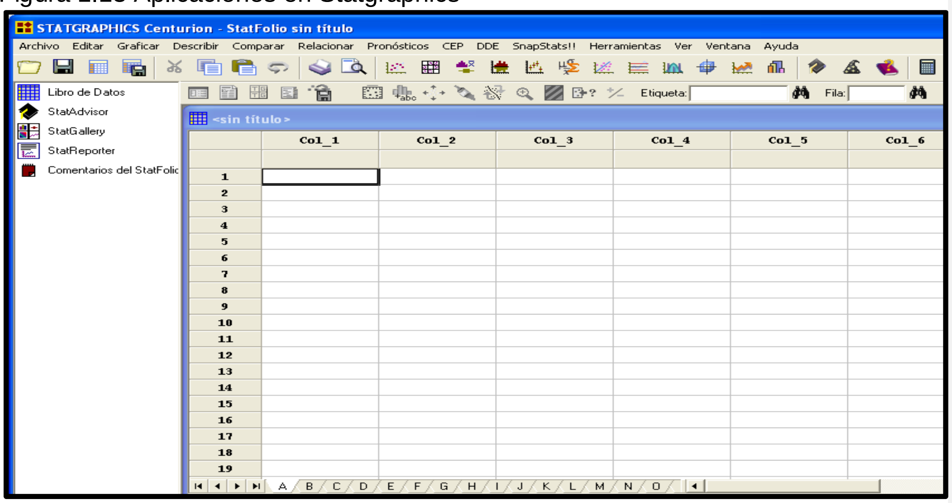
Con formato: Centrado

1.6 STATGRAPHICS CENTURIÓN VERSIÓN XVI

Para resolver los ejercicios y casos planteados en este texto, se utilizará el software Statgraphics centurión versión XVI.

Cuando se ingresa al programa Statgraphics centurión versión XVI, se despliega la ventana que se ilustra en la siguiente figura.

Figura 1.13 Aplicaciones en Statgraphics



Fuente: Los Autores

Menú de STATGRAPHICS centurión versión xvi.

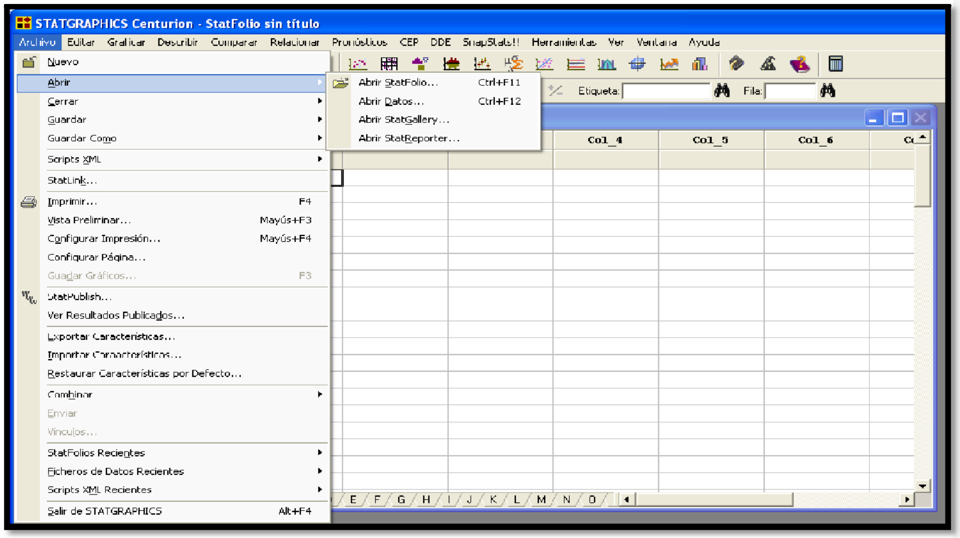
- **Archivo:** Incluye funciones como abrir, cerrar, guardar, imprimir (vista previa de impresión, configuración de impresión, configuración de pagina), combinar y publicar en internet un StatFolio, datos, Statgallery y StatReporter.

  StatFolio: Son archivos de datos de Statgraphics u otra aplicación.

  Statgallery: Es una ventana en la que se pueden pegar las graficas que son creadas en Statgraphics, la galería son varias ventanas en la cual cada una contiene una grafica.

StatReporter: Es una ventana donde se pueden crear y guardar reportes como archivos. Además de copiar y pegar tablas y graficas creadas en algún otro lugar usando las opciones copiar y pegar en el menú editar.

Figura 1.14 Aplicaciones en Statgraphics



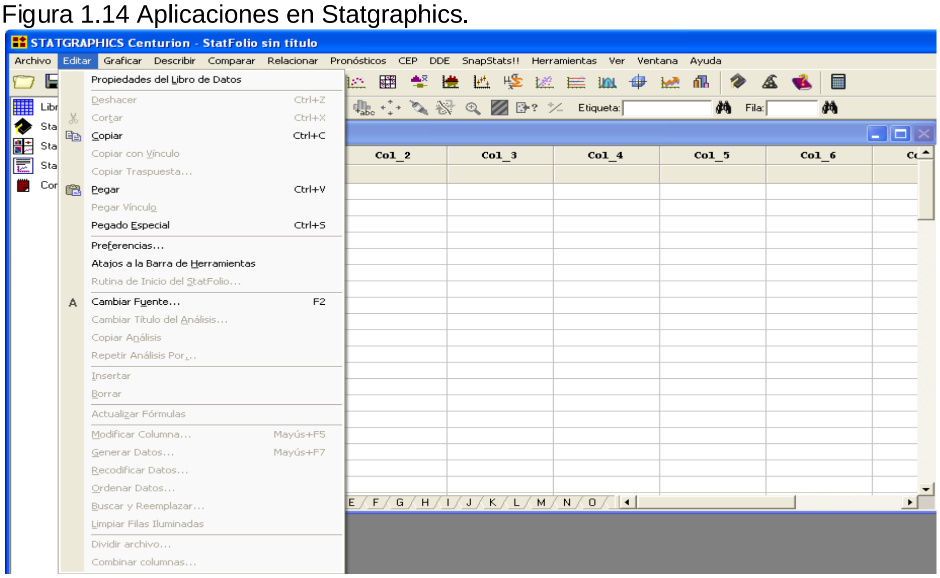
Fuente: Los Autores

- **Editar:** Es una opción que ofrece la siguientes funciones:
  - Fijar las propiedades de cada fuente de datos en un libro de datos.
  - Deshacer la última operación que se realice en el libro de datos.
  - Cortar un objeto o texto seleccionado.
  - Copiar en el portapapeles de Windows los datos seleccionados, o la tabla o gráfica seleccionada.
  - Fijar preferencias para la operación del sistema y procedimientos estadísticos comunes.
  - Definir un conjunto de operaciones a ejecutar cuando se carga un StatFolio (Rutina de Inicio del StatFolio)

    Cambiar la fuente que se usa en la ventana activa.

- Cambiar el título de una ventana de análisis.
- Crear una copia de la ventana de análisis resaltada.
- Insertar renglones (filas), columnas o celdas en una hoja de datos.
- Borrar los datos u objeto(s) resaltados.
- Actualizar Fórmulas en la hoja de datos activa.
- Cambiar el nombre, comentario o tipo de una columna de datos.
- Crear y añadir datos a una columna de la hoja de datos usando una expresión.
- Recodificar los datos en una columna de la hoja de datos.
- Ordenar los datos en una hoja de datos.

Figura 1.14 Aplicaciones en Statgraphics.



Fuente: Los Autores

- **Graficar:** Sirve para crear gráficos estadísticos entre los cuales tenemos:

Figura 1.15 Aplicaciones en Statgraphics

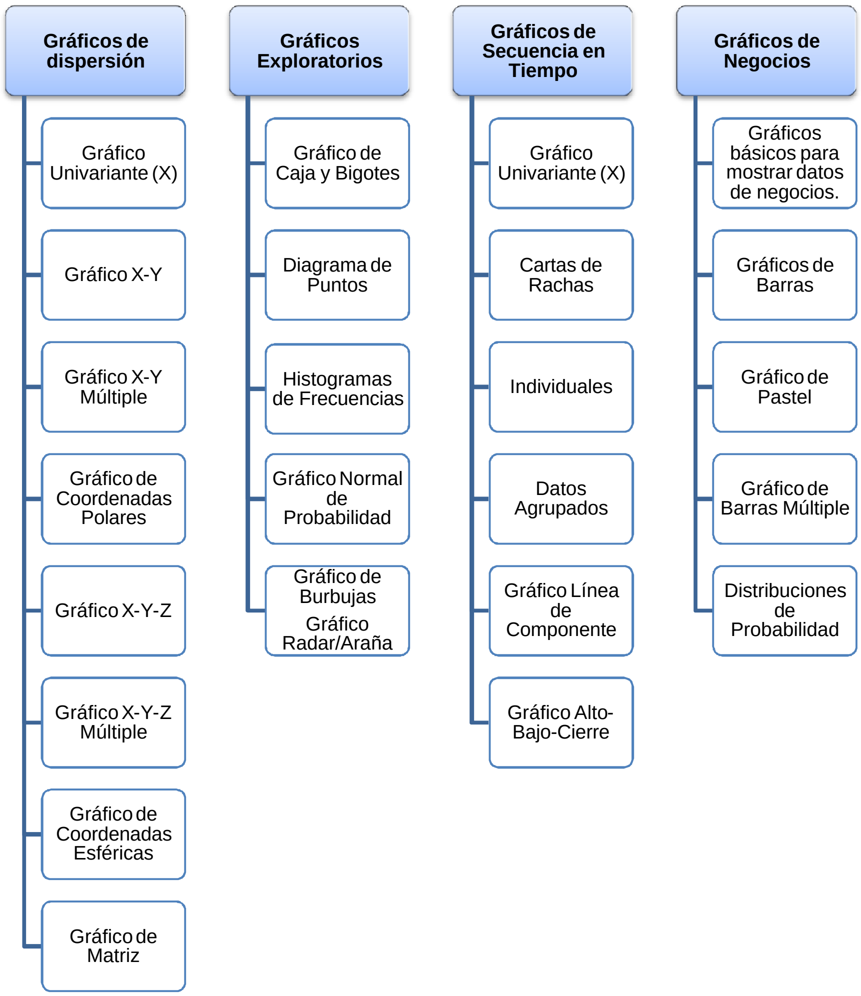
Fuente: Los Autores

- **Describir:** Sirve para aplicar procedimientos estadísticos, tales como:

Figura 1.16 Aplicaciones en Statgraphics

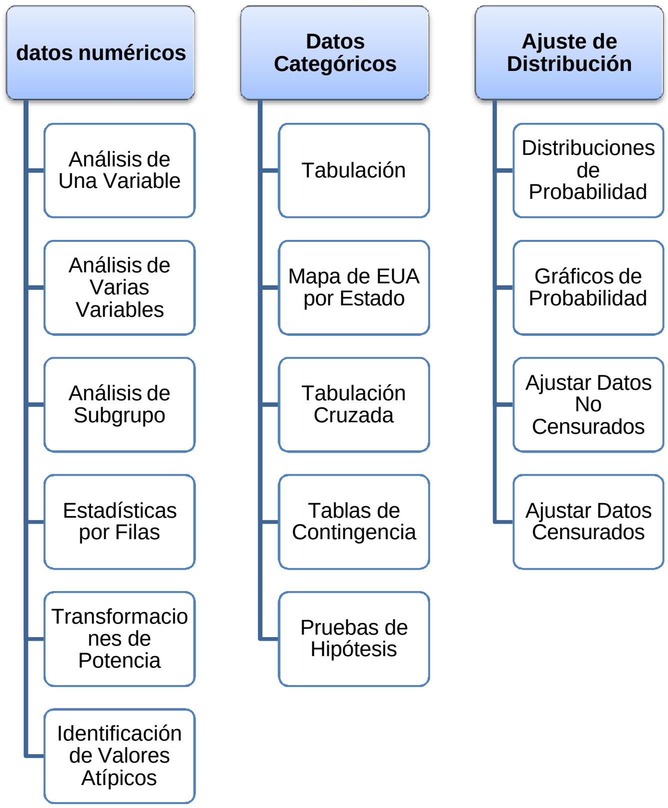
Fuente: Los Autores

Diagrama 1.1. Funciones de Gráficos



Fuente: Los Autores

Diagrama 1.2. Funciones para ingresar Datos



Fuente: Los Autores.

- **Comparar**: incluye procedimientos estadísticos para comparar dos o más grupos de datos, incluyendo procedimientos de dos muestras y de ANOVA.

Figura 1.17 Aplicaciones en Statgraphics

Fuente: Los Autores

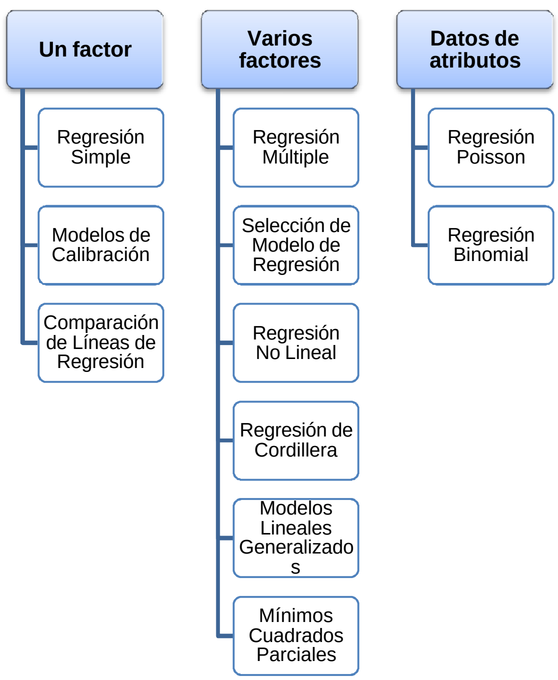
- A través de esta opción se pueden realizar análisis de:
- Dos Muestras
- Muestras Independientes
- Muestras en Pares
- Pruebas de Hipótesis
- Varias Muestras
- Comparación de Varias Muestras
- Comparación de Proporciones
- Comparación de Tasas
- Análisis de Varianza
- Gráfico de Medias de Factor
- ANOVA Simple
- ANOVA Multifactorial
- Pulido de Mediana de Tabla Bidireccional
- Componentes de la Varianza
- Modelos Lineales Generalizados

• **Relacionar:** procedimientos estadísticos para ajustar modelos de regresión.

Figura 1.18 Aplicaciones en Statgraphics

Fuente: Los Autores.

Diagrama 1.3. Funciones para Relacionar



Fuente: Los Autores

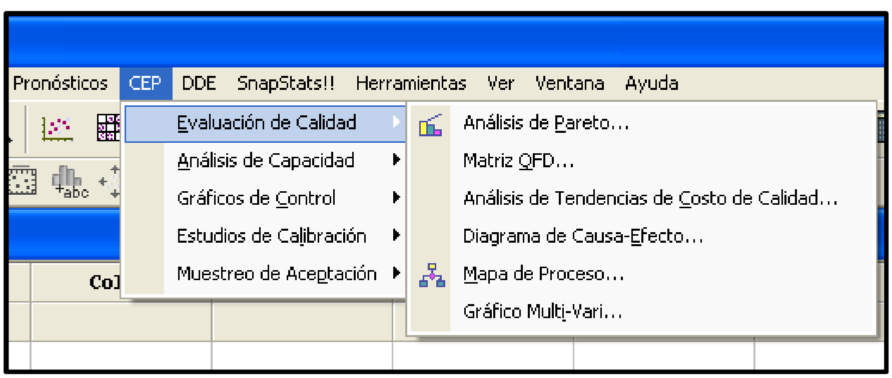
- **Pronosticar**: Permite realizar pronósticos de cambios en los procesos a través del tiempo para tomar la toma de decisiones. Los pronósticos de pueden hacer mediante un modelo definido por el usuario o pronósticos automáticos.

Figura 1.19 Aplicaciones en Statgraphics

Fuente: Los Autores

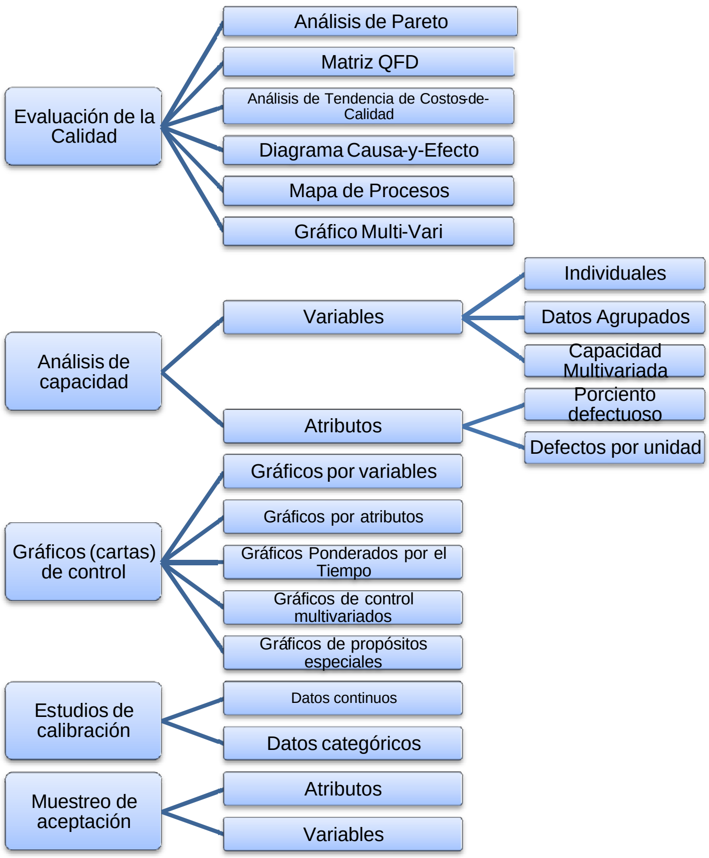
- **CEP:** Denotan las siglas de control estadístico de procesos. Esta herramienta emplea las técnicas para el control estadístico de la calidad.

Figura 1.20 Aplicaciones en Statgraphics



Fuente: Los Autores

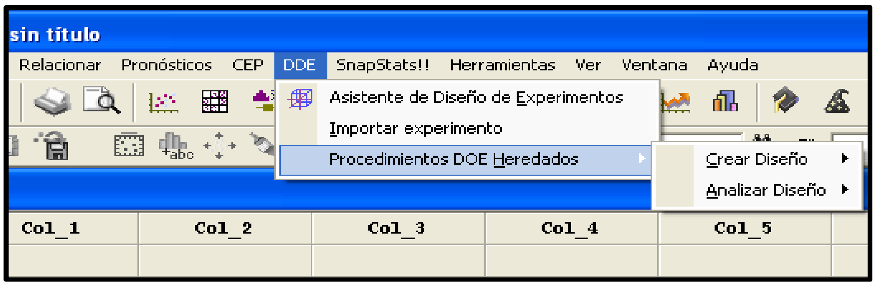
Diagrama 1.4. Funciones CEP



Fuente: Los Autores

- **DDE:** Sirve para crear diseños de experimentos, incluye las siguientes funciones para el análisis de experimentos:
  - Asistente de diseño de experimento
  - Importar experimento
  - Procedimientos de diseño de experimento

Figura 1.21 Aplicaciones en Statgraphics



Fuente: Los Autores.

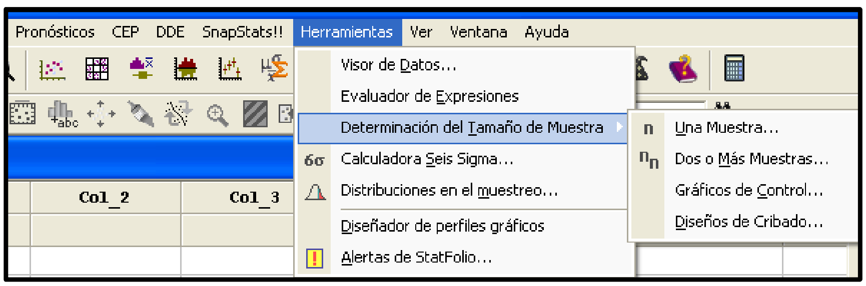
- **SnapStats!!:** Incluye los siguientes procedimientos estadísticos:
    - Análisis de una muestra
    - Comparación de dos muestras
    - Comparación de muestras pareadas
    - Comparación de varias muestras
    - Ajuste de curva
    - Evaluación de capacidad para datos individuales y agrupados
    - Estudios R&R
    - Pronósticos automáticos

Figura 1.22 Aplicaciones en Statgraphics

Fuente: Los Autores.

- **Herramientas:** Incluye procedimientos estadísticos, tales como:
  - Visor de datos
  - Evaluador de expresiones
  - Determinación del tamaño de muestra
  - Calculadora seis sigma
  - Distribuciones en el muestreo
  - Diseñador de perfiles gráficos
  - Alertas de StatFolio

Figura 1.23 Aplicaciones en Statgraphics



Fuente: Los Autores.

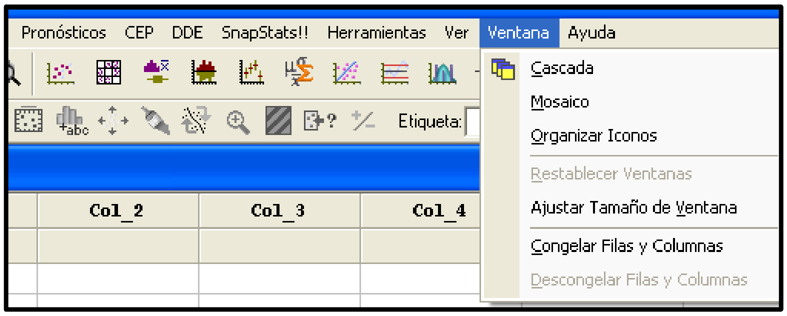
- **Ver:** Incluye opciones básicas para visualizar la barra de herramienta, la barra de estado, barra de navegación, la barra de tareas y el StatAdvisor.

Figura 1.24 Aplicaciones en Statgraphics

Fuente: Los Autores.

- **Ventana:** Sirve para listar de ventanas abiertas.

Figura 1.25 Aplicaciones en Statgraphics



Fuente: Los Autores.

- **Ayuda:** Permite el acceso a los manuales de ayuda y al sitio Web de STATGRAPHICS.

Figura 1.26 Aplicaciones en Statgraphics

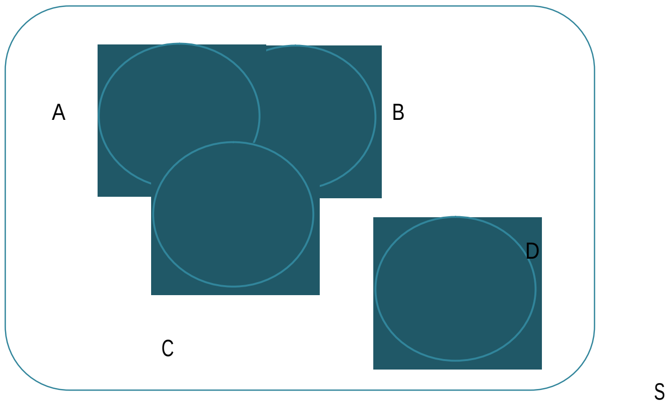
Fuente: Los Autores.

EJERCICIOS PROPUESTOS

CAPÍTULO N° 1.

Ejercicios Propuestos

1. En el siguiente diagrama de Venn se muestran los conjuntos: A, B, C y D.



Sombree la región que corresponda a cada uno de los siguientes eventos:

a. (• • •) • •

b. D’

c. (• • •)’

d. (• • •)’ • •

2. En la cafetería de la Universidad Tecnológica de Bolívar, se evalúan tres eventos.

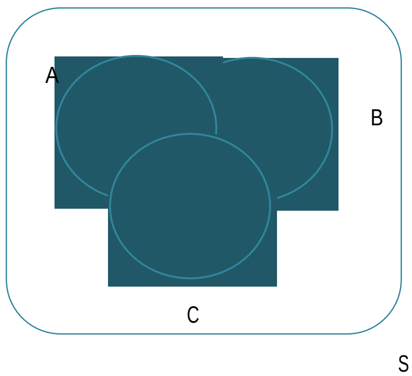
A: El estudiante encuentre la chocolatina del sabor deseado

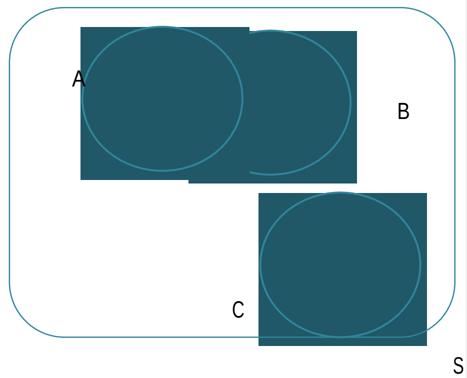
B: El estudiante encuentre la chocolatina del tamaño deseado

C: El estudiante encuentre la chocolatina con la dureza que desea

Esto se hace con el fin de conocer si el estudiante, como cliente de la cafetería queda satisfecho al encontrar la chocolatina con las características deseadas.

Se le pide a usted que interprete los siguientes diagramas de Venn:





3. El proceso de fabricación de un lapicero se ensamblan tres piezas A, B, y C. Si todas las piezas son ensambladas correctamente se puede vender el lapicero a los clientes, debido a que cumple con las especificaciones de calidad, de lo contrario se reprocesa. Muestre en diagramas de Venn, las dos situaciones.

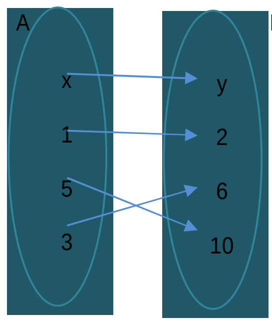
4. En los números naturales existe un orden de conteo, el hecho de que existan números mayores que o menores dentro de los mismos, se le conoce como:

a. Desigualdad
b. Diferencia
c. Igualdad
d. Mezcla

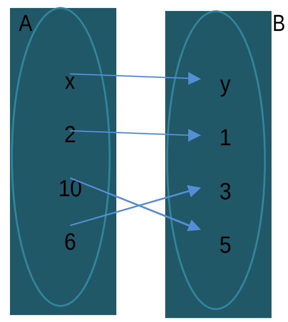
5. Juan tiene dos hermanos, llamados Julián y María, él quiere regarles unos dulces que ha comprado, pero quiere que sus dos hermanos queden satisfechos, debido que Julián come mucho ha decidido darle dos dulces a Julián por cada dulce que le de a María.

Sea A, el conjunto de los dulces que Juan le regala a María y B el conjunto de dulces que Juan le regala a Julián. Escoge el diagrama que muestra esta relación de conjuntos.

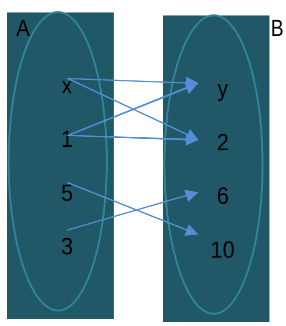
a.



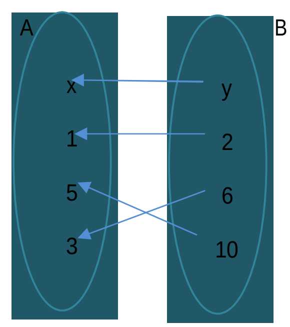
b.



c.



d.

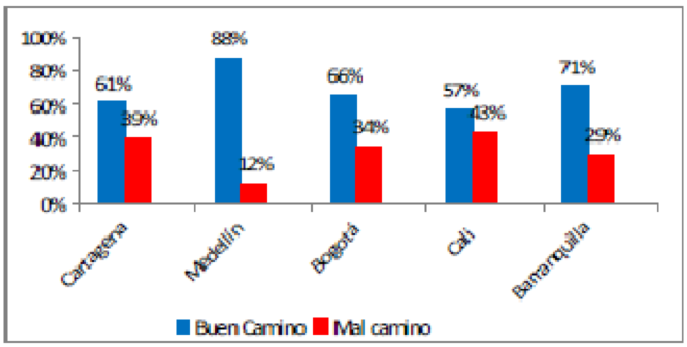


6. Del ejercicio anterior:

a. Representa a través de una función matemática la relación de los conjuntos A y B.
b. Menciona cuál de las graficas no cumple con alguna de las propiedades de asociación de conjuntos. Argumenta tu respuesta.

7. Siguiente grafica muestra los resultados de una encuesta relacionada con la percepción ciudadana sobre la calidad de vida en cinco ciudades colombianas, realizada en diciembre de 2008 en el marco del proyecto Cartagena Como Vamos.

Grafico 1.6 Estado del carro y del camino



Fuente: Proyecto Cartagena Como Vamos

De la gráfica anterior, se puede concluir:

a. El nivel de optimismo en Cali es mayor que el nivel de optimismo en Cartagena
b. El nivel de optimismo en Barranquilla es menor que el nivel de optimismo en Cartagena
c. El nivel de optimismo en Cartagena es menor que el nivel de optimismo en Medellín
d. El nivel de optimismo en Bogotá es mayor que el nivel de optimismo en Medellín

8. El siguiente cuadro muestra la distribución porcentual de las principales importaciones registradas en Colombia, según país de origen, de acuerdo al boletín de prensa del Comportamiento de las Importaciones y Balanza Comercial (Comercio Exterior) en diciembre de 2009, emitido por el

Departamento de Impuestos y Aduanas Nacionales (DIAN) el 16 de febrero de 2010.

Tabla 1.5. Distribución porcentual de las principales importaciones, según el país de origen de Enero – Diciembre (2009-2008)

País	Millones de dólares CIF	Part (%)	País	Millones de dólares CIF	Part (%)
	2009			2008	
Estados Unidos	9 456	28,7	Estados Unidos	11 437	28,8
China	3 715	11,3	China	4 549	11,5
México	2 298	7,0	México	3 126	7,9
Brasil	2 147	6,5	Brasil	2 328	5,9
Francia	1 457	4,4	Alemania	1 557	3,9
Alemania	1 338	4,1	Venezuela	1 198	3,0
Argentina	1 043	3,2	Japón	1 153	2,9
Japón	825	2,5	Argentina	920	2,3
Ecuador	695	2,1	Corea	920	2,3
Corea	680	2,1	Francia	884	2,2
Canadá	675	2,1	Ecuador	810	2,0
Perú	623	1,9	Canadá	795	2,0
Demás países	7 946	24,2	Demás países	9 991	25,2

Fuente: DIAN Cálculos DANE

Del anterior cuadro calcula:

a. La sumatoria de la participación (%) de las importaciones de enero a diciembre de 2009.
b. Suponiendo que, el Departamento de Impuestos y Aduanas Nacionales (DIAN) esperaba que las importaciones para el año 2009 incrementara la participación en 0,1 de cada país, de acuerdo a las participaciones en las importaciones registradas en el año 2008. Calcula la sumatoria de la participación esperada para el año 2009.

9. Escribe F, si la afirmación es falsa o V si es verdadera.

a. Los números irracionales están conformados por los números enteros negativos y los números racionales ( )

b. Los números racionales están conformados por los números enteros negativos y los números racionales ( )
c. La unidad de la varianza de una variable es la unidad de la variable al cuadrado ( )
d. La sumatoria •⁚..(•. + •) no es igual a •⁚.. •. + •• ( )
e. Cualquier número real puede escribirse como un número decimal ( )

10. Una familia de estrato 4 de la ciudad de Cartagena, tiene un consumo de energía en Kw/h, se conoce el costo de 1Kw/h, además en la facturación se relaciona el costo del alumbrado público que es tres cuartas partes mayor que el subsidio, y el subsidio a estratos 0 y 1 que es 0.2% del costo total de la factura. Plante la función con la cual se puede determinar el costo total de la factura.

ESTUDIO DE CASOS

CASO 1

En la empresa “Lechecita y Algo Mas”, se producen productos derivados de la leche, como son: mantequilla, quesos, leches con sabores, kumis, suero, y yogurt.

El yogurt es un producto lácteo fermentado, levemente ácido, de cultivo semisólido que es producido por homogenización y pasteurización.

La leche entera o descremada, fresca o en polvo, puede ser utilizada para producir yogurt mediante el siguiente proceso:

Inicialmente, la leche es bombeada a un tanque de almacenamiento. Luego es bombeada a través de un filtro hacia el clarificador que va a utilizar una fuerza centrífuga para obtener algunas impurezas insolubles en la leche. En seguida la leche es pasteurizada y desinfectada en un sistema de pasteurización de temperaturas muy altas (UHT) el cual disminuye el crecimiento de las bacterias de ácido láctico a un nivel aceptable; después, los cultivos de steptococus termofílica y el lactobacilo bulgaris, que crecen en un ambiente controlado son injertados dentro de la leche esterilizada y luego pasan a los tanques de fermentación donde son fermentados a una temperatura de 40°C por 30 horas. Luego, se mezclan los saborizantes y algunos aditivos para ser esterilizados en un sistema de esterilización de altas temperaturas y corto tiempo (HTST).

El yogurt fermentado y los aditivos esterilizados son mezclados y homogenizados. Seguidamente, el yogurt es colocado en cajas de polipropileno o polietileno por una máquina rellenadora, finalmente los envases de yogurt se colocados en el almacén de refrigeración.

El gerente de la empresa desea saber cómo sería la interrelación entre cada uno de los procesos que se llevan a cabo para la producción del yogurt. Además saber cuáles son los procesos llamados independientes, de igual manera el proceso que se relaciona con todos los procesos.

CASO 2

En un estudio realizado por el Departamento Administrativo Nacional de Estadística (DANE) acerca de los principales indicadores del mercado laboral relacionado con la seguridad social. Se tomo como referencia la población de personas que se encuentran laborando y están afiliadas en seguridad social, salud, y pensión; de dicha población se tuvo en cuenta muestras en cada uno de los trimestres de tipo móvil desde Diciembre de 2008 hasta Enero de 2010.

Se utilizo un mecanismo de encuestas aleatorias, para ser aplicadas en la muestra teniendo en cuenta la población.

Los datos obtenidos se muestran en la siguiente tabla:

Concepto	2009								2010			
	Dic-feb	Ene-Mar	Feb-Abr	Mar-May	Abr-Jun	May-Jul	Jun-Ago	Jul-Sep	Ago-Oct	Sep-Nov	Oct-Dic	Nov09-Ene 10
Total Ocupados Nacional	17493	17650	18103	18375	18491	18489	18396	18321	18655	18974	19248	18867
Salud	15280	15347	15723	15970	16129	16178	16157	16153	16451	16778	17033	16759
R. Contributivo.	7861	7856	8157	8338	8351	8373	8302	8343	8403	8479	8488	8313
R. Especial	377	389	398	413	432	425	405	393	385	415	421	440
R. Subsidiado	7039	7101	7167	7217	7344	7377	7447	7416	7662	7884	8121	8003
No sabe	2	2	2	2	3	2	3	2	1	1	2	3
Pensiones	5271	5288	5455	5550	5573	5595	5550	5645	5582	5602	5512	5452

a) Realice un informe ejecutivo donde muestre los resultados obtenidos a partir de la encuesta realizada, comparando los trimestres y cuál de los semestres se encuentren mayores personas afiliadas, además saque conclusiones de porque se presenta el caso de tener mayor número de personas afiliadas en salud que en pensión, en este informe se debe mostrar en cual trimestre se presenta mayor y el menor aporte entre los régimen y determinar en este tiempo los hechos históricos que ayudaron a que se diera esta situación

b) Suponiendo que exista una relación entre la proporción de afiliados a salud y los afiliados a pensión, establezca cuanto es el número de personas que se encuentran afiliadas a pensión para cada régimen de salud.

CAPÍTULO 2

CONCEPTOS BÁSICOS DE ESTADÍSTICA.

INTRODUCCION

La estadística está presente en la lógica humana; el razonamiento de situaciones con datos cualitativos y/o cuantitativos da como resultado teorías que al paso del tiempo pueden servir como datos para futuros escenarios, permitiendo al ser humano predecir resultados, así, la estadística puede remontarse a datos pasados o presentes para estudiarlos y generar nuevas interpretaciones del acontecimiento presente o prever el futuro, es por esto que la estadística es definida como "La ciencia que se ocupa de los métodos y procedimientos para recoger, clasificar, resumir y analizar los *datos*, siempre y cuando la variabilidad e *incertidumbre* sea una causa intrínseca de los mismos; así como de realizar *inferencias* a partir de ellos, con la finalidad de ayudar a la toma de *decisiones* y en su caso formular *predicciones*"

La estadística actualmente es una herramienta importante y fundamental en todos los aspectos de la sociedad, y su evolución constante como ciencia facilita el análisis de la información y la solución de problemas reales, ayudando en la investigación en los diferentes niveles de la ciencia, convirtiéndose en uno de los pilares de ésta.

En este capítulo se incluirá los conceptos básicos en la estadística tales como la población, muestra, variables, parámetro, estadísticos que sirven como base para entender la estadística descriptiva, la cual se relaciona con el desarrollo y uso de técnicas para la cuidadosa recolección y efectiva presentación de la información cuyo fin es describir las características principales de los datos reunidos de tal modo que permitan sugerir o aventurar cuestiones a analizar en mayor profundidad.

Además se presentarán tres tipos de medidas de resumen: medidas de tendencia central, medidas de dispersión o variabilidad de los datos y medidas de ubicación, las cuales hacen parte de los conceptos fundamentales de la estadística ya que muestran el comportamiento de los datos y son usadas en otras operaciones más complejas utilizadas para el análisis de datos.

COMPETENCIAS:

- Interpretacion de nociones básicas relacionadas con el manejo de información como población, muestra, variable aleatoria, distribución de frecuencias, parámetros y estadígrafos).
- Utilizacion de medidas de tendencia central (media, mediana, moda) para interpretar comportamiento de un con- junto de datos.
- Utilizacion de modelos (diagramas de árbol, por ejemplo) para discutir y predecir posibilidad de ocurrencia de un evento.
- Interpretacion y utiliacion de conceptos de media, mediana y moda y explicito sus diferencias en distribuciones de distinta dispersión y asimetría.
- Seleccion y uso de algunos métodos estadísticos adecuados al tipo de problema, de información y al nivel de la escala en la que esta se representa (nominal, ordinal, de intervalo o de razón).

2. CONCEPTOS FUNDAMENTALES DE ESTADISTICA

Desde los comienzos de la civilización, la estadística ha estado presente de una u otra manera en formas sencillas, pues se utilizaban representaciones gráficas y otros símbolos sobre pieles y paredes de cuevas para representar animales, personas o ciertas cosas.

La mayor parte de las primeras aplicaciones de la estadística consistieron principalmente en la presentación de datos en forma de tablas y gráficas. Este campo se desarrolló con rapidez llegando a incluir gran variedad de métodos para ordenar, resumir y expresar en alguna forma las características de un conjunto de números.

Durante el siglo XVIII empieza el auge de la estadística descriptiva en asuntos sociales y económicos, y es a finales de ese siglo y comienzos del XIX cuando se comienzan a asentar verdaderamente las bases teóricas de la teoría de probabilidades con los trabajos de Joseph Louis Lagrange y Pierre Simon de Laplace, Carl Friedrich Gauss, y de Simeón-Denis Poisson. Previamente, cabe destacar el descubrimiento de la distribución normal por Abraham de Moivre, distribución que será posteriormente “redescubierta” por Gauss y Poisson.

En nuestros días, la estadística se ha convertido en un método efectivo para describir con exactitud los valores de los datos económicos, políticos, sociales, psicológicos, biológicos y físicos, y sirve como herramienta para relacionar y analizar dichos datos. Sin embargo, la estadística implica más que reunir y tabular los datos, sino sobre todo en interpretar esa información.

2.1 DEFINICION DE LA ESTADISTICA

La estadística es “La ciencia que se ocupa de los métodos y procedimientos para recoger, clasificar, resumir y analizar los *datos*, siempre y cuando la variabilidad e *incertidumbre* sea una causa intrínseca de los mismos; así como de realizar *inferencias* a partir de ellos, con la finalidad de ayudar a la toma de *decisiones* y en su caso formular *predicciones*”[4]

La estadística se clasifica en:

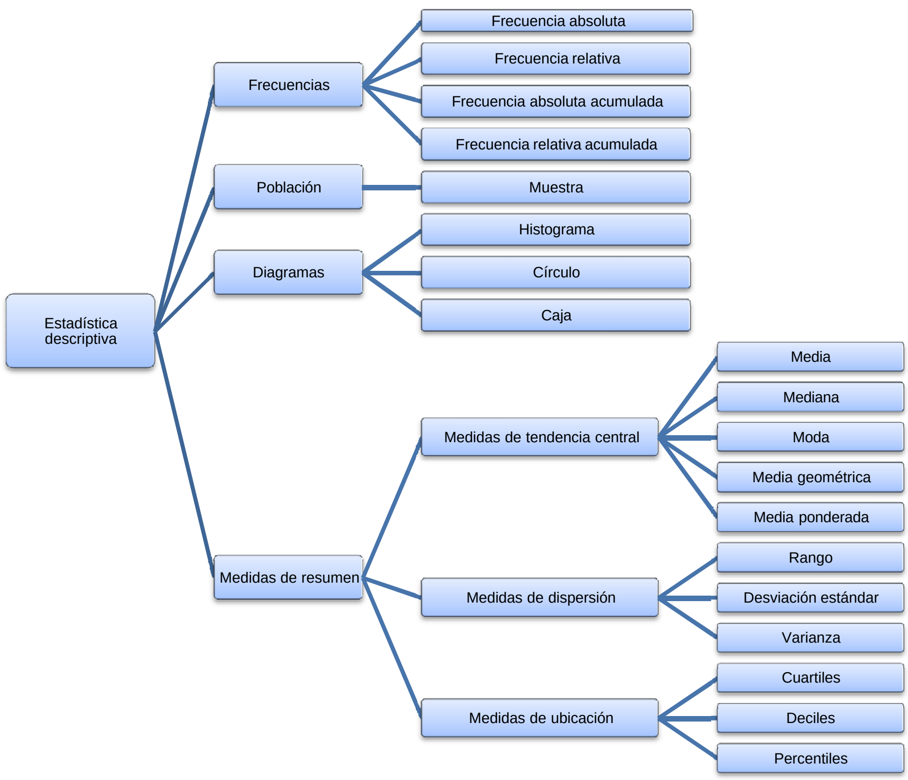
- ***Estadística descriptiva*** es la rama o disciplina que se relaciona con el desarrollo y uso de técnicas para la cuidadosa recolección y efectiva presentación de la información la cual tiene como objetivo describir las características principales de los datos reunidos. Básicamente hay tres tipos de medidas de resumen: medidas de tendencia central, medidas de dispersión o variabilidad de los datos y medidas de ubicación.

  Uno de los objetivos de la Estadística Descriptiva es presentar los datos de tal modo que permitan sugerir o aventurar cuestiones a analizar en mayor profundidad.

- ***Estadística Inferencial:*** Sólo se centra en tomar una pequeña muestra representativa de la población y a partir de la información de la misma, infiere que el resto de los elementos de la población tienen el mismo comportamiento. En caso de que un muestreo para cierto estudio no sea factible realizarlo por cuestiones de tiempo, recursos o costo; se puede calcular un tamaño de muestra para medir solo algunos elementos de la población; posteriormente se infiere que el resto de los elementos de la población se comportan igual que la muestra tomada.

[4] Tomado de http://www.bioestadistica.uma.es/libro/node3.htm

Diagrama 2.1 Estadística Descriptiva y sus componentes



Fuente: Los Autores

2.2 POBLACIÓN (•):

La población es el conjunto de personas, datos u objetos que poseen características comunes observables en un espacio y tiempo determinado.

En el proceso investigativo, la población corresponde al conjunto de referencia sobre el cual se va a desarrollar la investigación o estudio.

En una investigación, cuando se selecciona una población para su estudio se debe tener en cuenta las siguientes características:

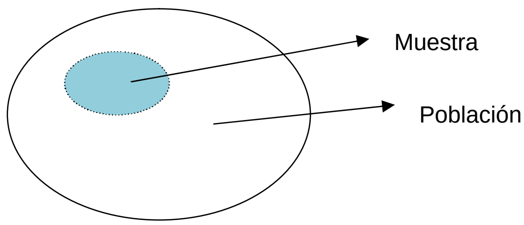
- **Homogeneidad:** se refiere a la similitud en las características de las variables a estudiar de la población, un ejemplo es la población de estudiantes de Ingeniería Industrial de la Universidad Tecnológica de Bolívar, por lo cual, todos los miembros que hacen parte de esta población necesariamente deben estar estudiando en la Universidad Tecnológica de Bolívar.
- **Cantidad:** Hace referencia al tamaño de la población. La población se puede clasificar de acuerdo a su tamaño en:

  Ø Población finita: Es el conjunto que incluye un número limitado de elementos, que se puede determinar a través de conteo. Ejemplo: El número de pacientes que se encuentran hospitalizados en la sala de urgencia del Hospital Metropolitano, la cantidad de silla instaladas en la primera sala de cine del Centro Comercial Plaza Caribe.

  Ø Población Infinita: Es el conjunto en el cual no existe limite en el número de elementos que lo conforman. Ejemplo: Los productos de belleza que hay en el Mercado, el conjunto de resultados obtenidos en la prueba ECAES realizada a los estudiantes de Psicología, la cantidad de taxis que transitan por la avenida Pedro Heredia el día Sábado.

2.3 MUESTRA (•):

La muestra es un subconjunto de la población.



Una muestra se selecciona con las siguientes finalidades:

- Conocer o inferir las características de toda población a partir del análisis y estudio de la muestra.
- Disminuir los costos de realización de la investigación, al realizar el estudio en una población numerosa se pueden elevar los costos.
- Disminuir la destrucción de objetos cuando se realiza el estudio, esto aplica a los procesos de ensayo destructivos.
- Acortar el tiempo de realización del estudio, debido que estudiar toda la población implica mayor tiempo que estudiar parte de la misma.
- Aumentar la calidad del estudio, al realizar las mediciones y observaciones a un número reducido de miembros de la población puede que sea más exactos que si se realizan al total de la población.
- La selección de muestras específicas pueden reducir la variabilidad en las características de la población al indicar criterios de inclusión o exclusión en el estudio.

Cuando se escoge una muestra se debe asegurar que esta sea confiable y representativa. Es decir, que la muestra tenga todas las características relevantes de la población.

Es importante destacar que la muestra no debe resultar sesgada, es decir, manipulada u orientada durante el proceso de selección, con el fin de establecer si la variable estudiada se comporta de acuerdo a una distribución de probabilidad.

2.4 VARIABLES:

Las variables son características que pueden ser medidas en el tiempo, para diferentes individuos u objetos de consideración, y de esta forma se obtienen los datos.

Matemáticamente, una variable es un símbolo que puede tomar valores de un conjunto de datos.

Ejemplo 2.1:

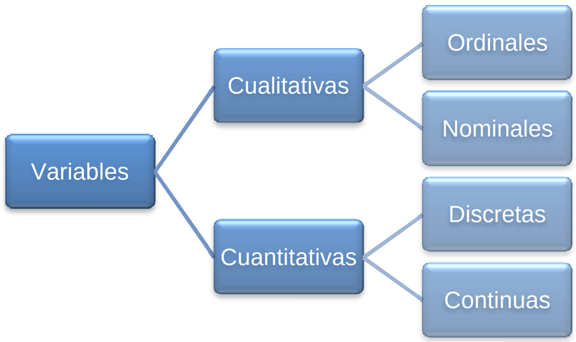
- = marca de zapatos que tiene puesto una persona
- = Cantidad de no conformidades detectadas en la auditoria
- = Distancia entre dos puntos A y B

2.3.1 Clasificación de las variables

Las variables pueden ser clasificadas como cuantitativas o cualitativas (categóricas), dependiendo si los valores presentados tienen o no un orden de magnitud natural (cuantitativas), o simplemente un atributo no sometido a cuantificación (cualitativa).

Un atributo corresponde a un valor específico e una variable, como ser el cpor ejemplo la variable estado civil puede ser: soltero, casado, viudo, unión libre.

Diagrama 2.2 Clasificación de las variables



Fuente: Los Autores

Variables Cualitativas: Son aquellas que indican cualidades o nombres. Puede ser dicotómica. Es decir, adopta un valor sin jerarquía entre sí tales como "si" o "no", "Femenino" o "masculino" y multicotómicas, si existe la posibilidad que adopten varios valores tales como grupo sanguíneo, religión, edad.

- **Variables nominales:** Son aquellas variables que no tiene un orden determinado. Por ejemplo: nombres de personas, raza, tipos de pizzas. Estas variables no tienen ningún orden inherentes a ellas ni un orden de jerarquía.

- **Variables ordinales:** Estas variables poseen un orden, secuencia o progresión natural esperable determinado, ejemplos: grados de escolaridad, nivel socioeconómico, meses del año. A pesar de este orden jerárquico no es posible obtener valoración numérica lógica entre dos valores.

Variables cuantitativas: Son aquellas variables cuyo valor son números.

- **Variables discretas:** Son aquellas que solo pueden adoptar un solo valor numérico, entero, con valores intermedios que carecen de sentido ejemplos: El número de estudiantes de cada género que tiene un salón de clase, El número de computadores que se encuentran en funcionamiento en una empresa.

- **Variables continuas:** Son aquellas que toman cualquier valor dentro de un rango numérico determinado, Ejemplos: La cantidad de minutos que gasta una persona en llenar 5 galones de agua, la estatura de los jóvenes que integran el equipo de baloncesto de la selección Bolívar, el valor del pasaje del Transcaribe.

Con formato: Sangría: Izquierda: 1,27 cm, Sin viñetas ni numeración

- Tanto las variables discretas como las continuas pueden agruparse costruyend intervalos, entre cuyos valores extremos se ubicaran las diferentes observaciones registradas. Sin embargo, solo las variables continuas pueden ser objeto de categorización mediante intervalos.

2.5 DATOS:

Un dato hace referencia a la información y puede definirse como un valor que puede tomar una variable; t~~T~~ambién se puede decir que es cuando una variable se mide en un conjunto de una unidad experimental.

2.5.1 Clasificación de datos estadísticos:

Los datos estadísticos se pueden clasificar en las siguientes formas:

- **Datos Cuantitativos:** Son todos aquellos datos que pueden medirse, cuantificarse o expresarse numéricamente. Ejemplo: Las edades de los estudiantes
- **Datos Cualitativos:** Son todos aquellos datos que pueden ser expresados por cualidades o atributos, Ejemplo: Las diferentes razas (Blanco, Moreno, mestizo, mulato, hindúes, etc.) de continente americano.

  La medición de estos tipos de datos se puede realizar:

- Nominalmente: Es la manera de categorizar los datos que no tiene una relación de orden entre sí, por ejemplo programas universitarios que ofrece la universidad, estado civil, etc.
- Ordinalmente: Son aquellos datos que presentan un orden o jerarquía entre las categorías. por ejemplo, totalidad de semestres a cursar de los programas tecnológicos.

2.6 PARÁMETRO

Definimos un parámetro como una medición numérica o valor representativo de la población, que describe algunas características de la misma, es decir, los parámetros sirven para caracterizar una población a través de un valor fijo (no aleatorio) generalmente desconocido, debido a la dificulad de observar todas y cada una de las unidades que conforman la población.ejemplos: El promedio de vida de los hombres colombianos; el promedio de horas extras dominicales laboradas por los trabajadores de una empresa; el porcentaje de desempleo de los jóvenes entre 20 y 35 años de la ciudad de Cartagena.

2.7 ESTADÍSTICO O ESTADÍGRAFO

Un estadístico es una medición numérica que describe algunas características de una muestra con el fin de comparar las características de una población o modelo estadístico[5]., es decir, los estadísticos surgen por la necesidad de estimar el valor de un parámetro y dicha estimación se realiza por medio de muestras, por tanto se puede definir como una variable aleatoria ya que va en función de las observaciones muestrales. ejemplos: La edad promedio de los primeros 10 estudiantes que ingresaron a estudiar ingeniería mecánica en la Universidad Tecnológica de Bolívar en el segundo semestre académico de 2009; 75 jóvenes

[5] TRIOLA, Mario F. “PROBABILIDAD ESTADÍSTICA”, Ed. Pearson Educación, Novena Edición Pag. 5

entre 15 y 19 años hacen parte del programa de planificación familiar de Profamilia, y de estos jóvenes el 20% utiliza métodos como el DIU (Dispositivo Intrauterino) para planificar

2.8 FRECUENCIA

Es el número de ocurrencias de un suceso en cierto periodo o espacio, ejemplos: Se desea evaluar las llegadas de las personas a una cola (un evento aleatorio) durante 30 minutos, de la observación realizada se obtuvo que durante estos treinta minutos llegaran 15 personas a la cola. Las 15 personas representan la frecuencia, es decir el número de personas que llegaron a la cola después de transcurridos 30 minutos.

2.8.1 Tipos de frecuencia:

- Frecuencia absoluta (•.): Es el número de veces que se repite el valor de una variable dentro del conjunto de datos.

  Las propiedades de las frecuencias absolutas son:

  a. La frecuencia absoluta de un conjunto de datos cualitativos o cuantitativos debe ser un número comprendido entre cero y el tamaño de la muestra (•), incluyendo 0 y •. Esto es: 0 • •. • •

  b. La sumatoria de la frecuencia absoluta de los datos cualitativos o cuantitativos es igual al tamaño de la muestra (n).

,

•
• •. = •
•••

- **Frecuencia relativa (•.):** Es el cociente entre la frecuencia absoluta (•.) y el tamaño de la muestra (n)

  Las propiedades de las frecuencias relativas son:

  a. La frecuencia relativa de un conjunto de datos cualitativos o cuantitativos debe ser un número comprendido entre cero y uno, incluyendo 0 y 1. Esto es: 0 • •. • 1

  b. La sumatoria de la frecuencia relativa de datos cualitativos o cuantitativos es igual a 1.

  •

  • •. = 1

  •••

  La frecuencia relativa puede expresarse como un porcentaje, ya que relacionan una parte del conjunto con el total.

Frecuencia absoluta acumulada (•.)**:** es la suma de las frecuencias absolutas de los valores menores o iguales que x_i. Se representa por F_i: $F_i + f_1 + f_2 + f_3 + ... + f_i$.

- **Frecuencia relativa acumulada** (•.)**:** es el cociente que se presenta entre la frecuencia absoluta acumulada (*Fi*)y el número total de datos, *N*. Es decir.

  •. = •. / •

2.8.2 CONSTRUCCIÓN DE TABLAS DE FRECUENCIA

Las tablas de frecuencias son un método para organizar y resumir datos indicando el número de veces que se repite cada valor de la variable estudiada.

Para construir la tabla de frecuencia se debe realizar los siguientes pasos:

1. Cuando los datos no están agrupados, se debe realizar el conteo del número de veces que se repite una variable
2. Para variables continuas, se debe determinar el valor máximo y mínimo dentro del conjunto de datos que se posean.
3. Para variables continuas, se calcula el rango del intervalo, esto es la diferencia entre el valor máximo o límite superior y el valor mínimo o limite inferior.
4. Para variables continuas, se debe calcular el numero de intervalos y la Marca de clase de cada intervalo
5. Finalmente se determina la frecuencia absoluta, la frecuencia relativa y las frecuencias absoluta acumulada y relativa acumulada.

Ejemplo 2.2:

En una empresa Metalmecánica se realizó un estudio para determinar la capacidad de una sustancia sólida para resistir deformación o abrasión de su superficie en un material (dureza del material), en este se registraron los siguientes datos:

3.2, 3.1, 2.8, 2.9, 3.3, 3.2, 3.1, 3.0, 3.1, 3.1, 2.7, 2.8, 2.9, 3.0, 3.2, 3.1, 3.1, 3.0, 3.0, 2.9, 2.9, 3.0, 3.0, 3.1, 3.0, 3.1, 3.4, 3.3, 3.3, 2.9, 2.9.

La unidad de medida de la dureza es el DIN

Primero se debe realizar conteo del valor de cada una de las variables

Tabla 2.1 Dureza del material.

Clase	•.	•.
1	2.7	1
2	2.8	2
3	2.9	6
4	3.0	7
5	3.1	8
6	3.2	3
7	3.3	3

Fuente: Los Autores

El segundo, tercer y cuarto paso no se realiza debido que la variable solo toma 7 clase de valores específicos y la amplitud del intervalo es pequeño.

Quinto paso: Hallar la frecuencia relativa (•.):

$$•_• = \frac{•_•}{•}$$

Para la primera clase:

$$•_• = \frac{•_•}{•} = \frac{•}{••} = •,•••$$

La formula anterior se aplica para cada una de las clases.

Luego, se puede expresar la frecuencia relativa en forma de porcentaje, esto es:

$$•_• = \frac{•_•}{•} \times •••\%$$

Para $•_• = •,••• \times •••\% = •,•\%$ y así sucesivamente.

Tabla 2.2 Tabla de Frecuencia del ejemplo 2.2

Clase	•.	•.	•.	•. (%)	•.	•.
1	2.7	1	0,033	3,3%	1	0,033
2	2.8	2	0,067	6,7%	3	0,100
3	2.9	6	0,200	20,0%	9	0,300
4	3.0	7	0,233	23,3%	16	0,533
5	3.1	8	0,267	26,7%	24	0,800
6	3.2	3	0,100	10,0%	27	0,900
7	3.3	3	0,100	10,0%	30	1,000
Total		30	1	100%	-	-

Fuente: Los Autores

2.9 LIMITES DE CLASE

Cada clase está limitada por dos límites; Limites Inferior y Limite Superior de clase.

El límite inferior se denotará por $\bullet_i$ y el límite superior por $\bullet_s$.

2.9.1 Intervalos De Clase

Son los intervalos en que pueden agruparse los datos de una variable continua.

Tipos de intervalos:

- Cerrados: Conjunto de todos los valores de la variable continua (•) que puede estar comprendido entre • y •, incluyendo • y •.
  Matemáticamente: • • • • •

Se denotan a través de llaves, así: [ ]

- Abiertos: Conjunto de todos los valores de la variable continua (•) que puede estar comprendido entre • y •, sin incluir • y •.
  Matemáticamente: • < • < •
  Se denotan a través de paréntesis, así: ( )
- Mixtos: Conjunto de todos los valores de la variable continua (•) que puede estar comprendido entre • y •. además incluye en valor de • o •.
  Matemáticamente: • < • • • y • • • < •
  Se denotan a través de paréntesis, así: ( ] y [ ) respectivamente.

2.10 AMPLITUD O RANGO DE CLASE (•)

La amplitud de una clase es la diferencia entre los límites superior e inferior. La amplitud se denota por •s – •i

2.11 MARCA DE CLASE (•.)

La marca de clase es el punto medio entre cada intervalo, y con este valor se pueden hallar muchos cálculos en los parámetros.

Ejemplo 2.3

Se realizó un segundo estudio de dureza en los materiales en la empresa Metalmecánica, con el fin de verificar si la clasificación de esta serie de datos era la correcta, a continuación se muestran los datos que se obtuvieron del estudio:

0.3, 1.5, 2.4, 2.8, 3.3, 3.5, 3.8, 4.2, 4.3, 3.8, 3.6, 3.4, 2.9, 2.5, 1.7, 0.7, 3.4, 3.6, 3.9, 4.4, 3.1, 2.6, 2.0, 1.1, 1.3, 2.2, 2.7, 4.7, 3.9, 3.7, 3.4, 3.2, 3.5, 2.8, 3.8, 4.1, 4.8, 1.5, 3.2, 1.3

La unidad de medida de la dureza es el DIN

Para construir la tabla de frecuencia se sigue los 5 pasos anteriormente descritos:

1. Hallar los límites inferior y superior de clase.

• . = 0.3 •••

•. = 4.8 •••

2. Rango del intervalo= **L_s – L_i =** 4.8 DIN - 0.3 DIN = 4.5 DIN• 5 DIN

   Después se divide 5DIN/0.5=10

   Este resultado aquí presente es el número de intervalos que se deben tener en cuenta para que el límite inferior de la clase pertenezca al intervalo, pero el límite superior no pertenece al intervalo, o viceversa, para la solución del ejemplo, se considero el intervalo mixto de la forma: • • • < •

   Calculando la frecuencia relativa, la frecuencia relativa acumulada y la frecuencia absoluta acumulada. Se tiene el resumen en la siguiente tabla:

Tabla 2.3 Tabla de Frecuencia de la dureza del material

$L_j – L_s$	Y_j	f_j	F_j	h_j	H_j
[0, 0.5)	0.25	1	1	0.025	0.025
[0.5, 1.0)	0.75	1	2	0.025	0.050
[1.0, 1.5)	1.25	3	5	0.075	0.125
[1.5, 2.0)	1.75	3	8	0.075	0.200
[2.0, 2.5)	2.25	3	11	0.075	0.2775
[2.5, 3.0)	2.75	6	17	0.150	0.425
[3.0, 3.5)	3.25	7	24	0.175	0.600
[3.5, 4.0)	3.75	10	34	0.250	0.850
[4.0, 4.5)	4.25	4	38	0.100	0.950
[4.5, 5.0)	4.75	2	40	0.050	1
	-	40	-	1	-

Fuente: Los Autores

2.12 REPRESENTACIÓN GRAFICA PARA LAS VARIABLES CUALITATIVAS

En el momento de llevar a cabo la representación de las variables cualitativas existen una serie de gráficos que son conocidos como los más comunes para representar este tipo de datos, estos son: el diagrama de barras y el grafico de sectores.

2.12.1 Grafico de Barras

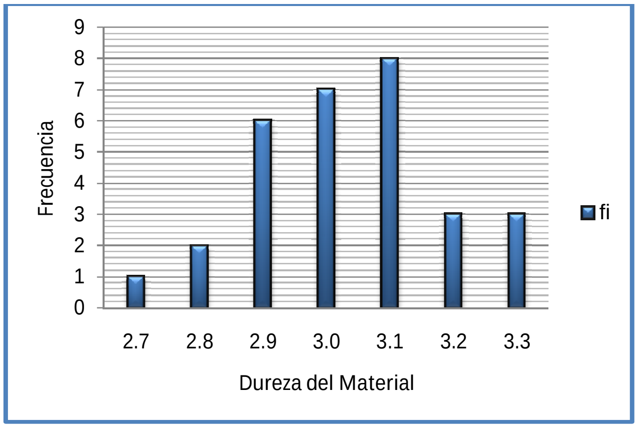
Es el que consiste en conjunto sucesivo de barras de la misma amplitud, con altura determinada por la frecuencia absoluta de cada una de las categorías en las que se encuentran las características establecidas, retomando el ejemplo 2.3 se desarrollara el siguiente grafico de barras:

Tabla 2.4 Dureza del Material

Clase	Dureza del material (DIN)	f_i
1	2.7	1
2	2.8	2
3	2.9	6
4	3.0	7
5	3.1	8
6	3.2	3
7	3.3	3
Sumatoria		30

Fuente: Los Autores

Grafico 2.1 Grafico de barra de dureza del material



Fuente: Los Autores

Ejemplo 2.4:

En las elecciones del Senado de la República de Colombia se obtuvieron los siguientes resultados:

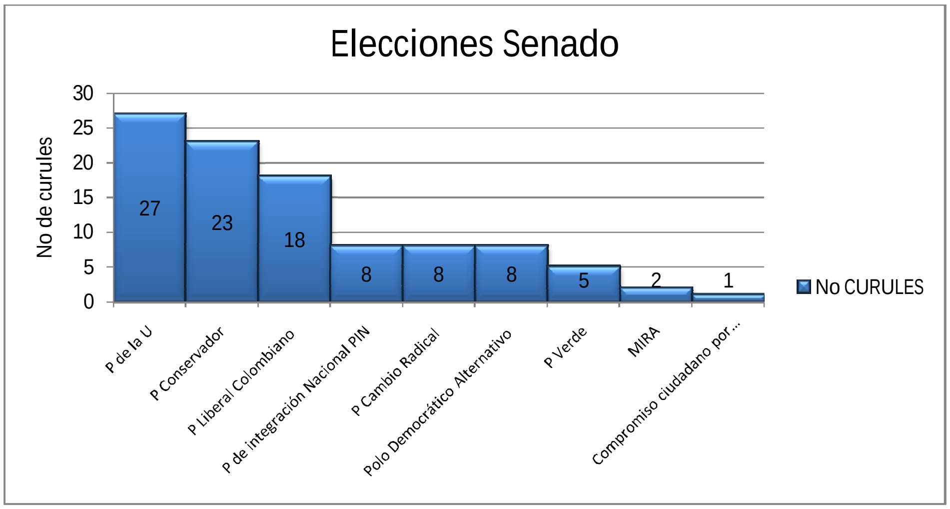
Tabla 2.5. Resultados de elecciones al Senado de Republica de Colombia 2010

NOMBRE LISTA	No CURULES
Partido de la U	27
Partido Conservador	23
Partido Liberal Colombiano	18
Partido de integración Nacional PIN	8
Partido Cambio Radical	8
Polo Democrático Alternativo	8
Partido Verde	5
MIRA	2
Compromiso ciudadano por Colombia	1

Fuente: http://www.registraduria.gov.co/elec2010/boletines.htm

Mediante el histograma podemos apreciar la distribución de las frecuencias.

Grafico 2.2 Histograma de Frecuencia



Fuente: Los Autores

2.12.2. Gráfico De Sectores ó Gráfico De Pastel

Consiste en un circulo dividido en sectores, cuyo tamaño es determinado por a frecuencia relativa de cada una de las categorías en la que se divide la variable. Retomando el ejemplo 2.4:

Para expresar el valor de la dureza en el material de forma porcentual, es necesario multiplicar por 100% el valor de la frecuencia relativa.
Otra forma de hallar el porcentaje es a partir de la siguiente fórmula:

Para el primer dato,

$$porcentaje\ (h_i) = \frac{valor\ de\ la\ variable \times 100\%}{sumatoria\ del\ total\ de\ datos}$$

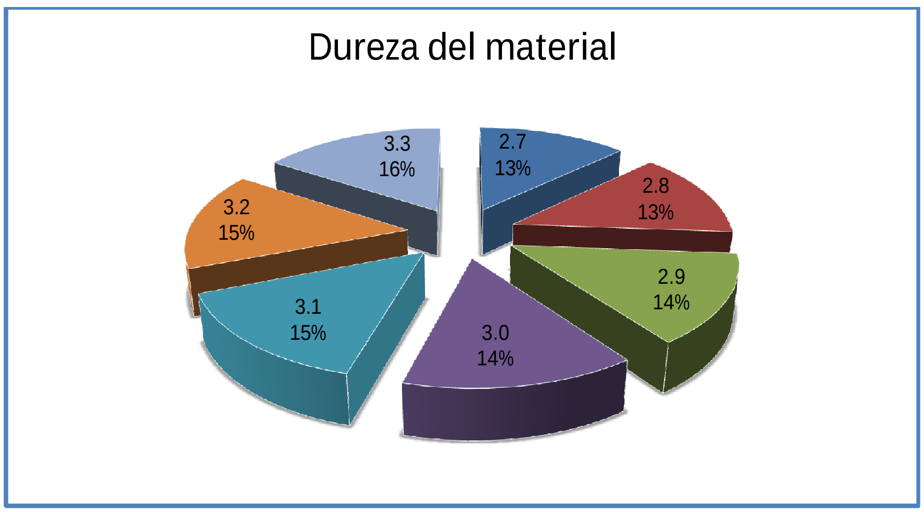
Y así sucesivamente para cada uno de los valores de la variable.

Tabla 2.6 Dureza en el material, frecuencia relativa, y porcentaje

Dureza en el material	**Frecuencia relativa**	**Porcentaje**
2.7	0.032	12,86%
2.8	0.065	13,33%
2.9	0.194	13,81%
3.0	0.226	14,29%
3.1	0.258	14,76%
3.2	0.097	15,24%
3.3	0.097	15,71%
Sumatoria	**1**	12,86%

Fuente: Los Autores

Grafico 2.3 Grafico Pastel de Dureza del material



Fuente: Los Autores

El Grafico anterior representa los sectores de la diferentes dureza del material donde se tiene que los materiales con mayor dureza ser encuentran ubicados en el rango de 3,3 DIN, y cada un cuenta con los diferentes grados de la dureza.

2.13 REPRESENTACIÓN GRAFICA DE LAS VARIABLES CUANTITATIVAS

En el caso de las variables cuantitativas ó datos agrupados en intervalos tenemos varias representaciones gráficas como: Histogramas, Polígonos de Frecuencias, Polígonos de Frecuencia acumulada, Curvas de Frecuencia, entre otros.

A continuación explicaremos cada una de las representaciones para este tipo de variable:

Histograma: Es la representación de una variable en forma de un grafico de barras, en donde las barras no guardan una separación entre si, este se obtiene construyendo cada intervalo de clase de la variable estadística con un rectángulo cuya área es proporcional a la Frecuencia de dicho intervalo. La suma de las áreas de los rectángulos debe ser uno, además si los intervalos tienen igual anchura la altura de los rectángulos es proporcional a las frecuencias de la clase

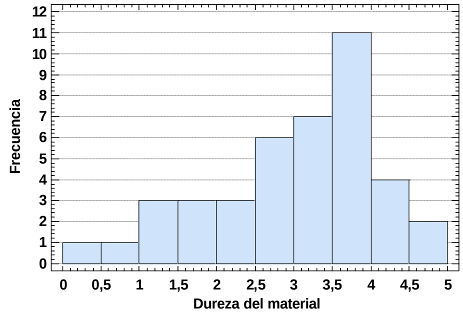
Ejemplo 2.5,

Tabla 2.7 Frecuencia de la dureza del material

•.	•,••	•,••	•,••	•,••	•,••	•,••	•,••	•,••	•,••	•,••	-
•.	1	1	3	3	3	6	7	10	4	2	40

Fuente: Los Autores

Grafico 2.4. Histograma de Frecuencia de la Dureza del Material

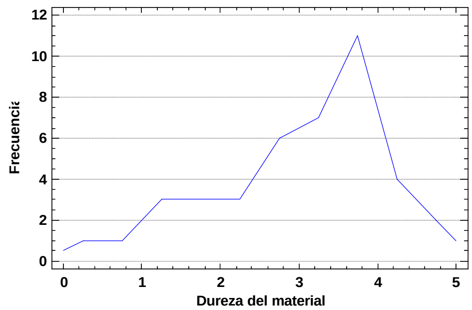


Fuente: Los Autores

Polígonos de Frecuencia: Este se construye uniendo con una línea poligonal los puntos, en otras palabras los puntos que forman las marcas de clase y sus correspondientes frecuencias absolutas, el polígono de frecuencia es la representación más real de los datos por cuanto que los valores individuales en cada intervalo son gradualmente más frecuentes a medida que se acercan al punto medio de la marca de clase.

Del ejemplo anterior podemos apreciar que el polígono de frecuencia sería:

Grafica 2.5 Polígono de Frecuencia de la Dureza del Material

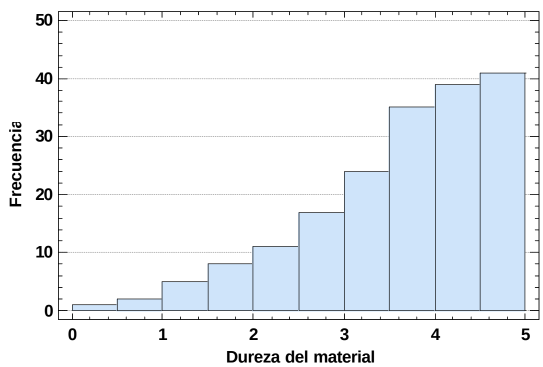


Fuente: Los Autores

Polígono de Frecuencias Acumuladas: Se construye tomando las frecuencias absolutas acumuladas hasta que cada uno de los limites de los intervalos y uniendo estos puntos son rectas. Esta representación grafica es importante ya que muchas veces el investigador enfoca su interés en el número de casos que caen por debajo o por encima de un valor específico más bien que el número de casos por cada intervalo.

Del ejemplo anterior podemos apreciar que el polígono de frecuencia sería:

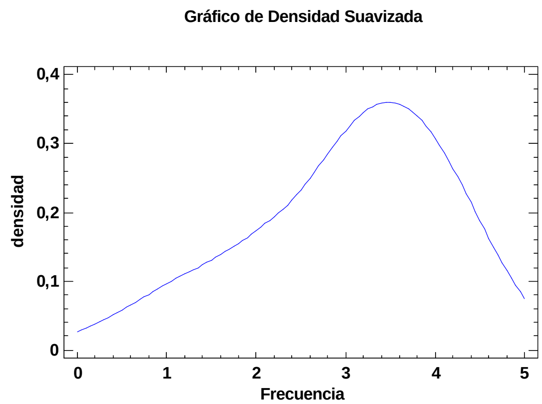
Grafica 2.6 Histograma de Frecuencia Acumulada de la Dureza del Material



Fuente: Los Autores

Curvas de Frecuencia: Este sugiere el uso de una curva suave como la representación idealizada de la distribución de la población.

Grafica 2.7 Curva de Frecuencia de la Dureza del Material



Fuente: Los Autores

2.14 MEDIDAS DE TENDENCIA CENTRAL

Existen tres medidas principales de tendencia central, las cuales son: la media aritmética ó media, mediana y moda

2.14.1 La media aritmética ó media:

La media aritmética, también llamada promedio o simplemente media es el valor que se obtiene al sumar todos los datos, y dividir el resultado entre el numero de valores de datos.

La media se denota por ••(se llama “x barra”), cuando se toman los valores de una muestra aleatoria este recibe el nombre de *media muestral.*

$$\bullet\bullet = \frac{\bullet\bullet\bullet\bullet\bullet\bullet\bullet}{\bullet}$$

Y se le llama • (mu minúscula), cuando se desea hallar la media de una población es de la siguiente forma:

$$\bullet\bullet = \frac{\bullet\,\dot{\bullet}\,\bullet\bullet\;\bullet\bullet}{\bullet}$$

Tabla 2.8 Notaciones de la Media

Notaciones
•Σ : sumatoria de un conjuno de valores •*x* : es la variable que representa cada uno de los datos individuales •*n* : es el número de valores de una muestra •N : es el número de valores de una poblacion

Fuente: Estadística y probabilidad de Ciro Martínez B.

Ejemplo 2.6:

En la escuela GRF está realizando una competencia de relevos entre los cursos del plantel educativo, y se han conformado grupos de diferentes tamaños, se toman 5 muestras las cuales tienen el siguiente número de estudiantes en la competencia:

Tabla 2.9. Estudiantes de GRF

	1	2	3	4	5
estudiantes	14	13	17	15	16

Fuente: Los Autores

Se desea calcular el número medio o media aritmética de la muestra de grupo de estudiantes es de:

$$\bullet\bullet = \frac{\bullet\,\dot{\bullet}\,\bullet\bullet\;\bullet\bullet}{\bullet}$$

$$= \frac{\bullet\bullet + \bullet\bullet + \bullet\bullet + \bullet\bullet + \bullet\bullet}{\bullet} = \frac{\bullet\bullet}{\bullet} = \bullet\bullet$$

• = •• Numero medio de estudiantes por grupo de la muestra

2.14.1.1 Cálculo de la media para Datos Agrupados

En algunas ocasiones se cuentan con datos en forma agrupada ó en forma de distribución de frecuencia, por eso se tiene otra forma para la estimación de la media y la cual se calcula; primero hallando el punto medio de cada clase, para lograr que los puntos medios se tengan en cifras cerradas se deben redondear las cantidades que se obtengan, seguidamente se multiplica cada punto medio por la frecuencia de observación de la clase, sumamos los resultados y dividimos entre el numero de observaciones de la muestra

La formula es la siguiente:

$$\bullet = \frac{\bullet_{\bullet\bullet}^{\bullet}(\bullet_{\bullet} \times \bullet_{\bullet})}{\bullet}$$

Tabla 2.10 Notaciones Media Muestral con datos agrupados

Notaciones
•= media de la muestra
• =Sumatoria de un conjunto de valores
•. =frecuencia del numero de observaciones de cada clase
•.= Punto medio de cada clase de la muestra
• = Numero de observaciones de la muestra

Fuente: Estadística y probabilidad de Ciro Martínez B.

Ejemplo 2.7:

La siguiente tabla de frecuencia muestra la distancia promedio recorrida por 385 niños de un plantel educativo

Tabla 2.11 Distancia promedio Recorrida

Clase (distancia mts)	Frecuencia (No. de Niños) (•.)
1 – 49.999	99
50 – 99.999	43
100 – 149.999	85
150 – 199.999	65
200 – 249.999	78
250 – 299.999	15
TOTAL	385

Fuente: Los Autores

Para calcular la media o distancia promedio que pueden recorrer los niños en metros se debe calcular la media para datos de frecuencia:

Tabla 2.11 Frecuencias de distancia promedio recorrida

Clase (distancia m)	(•.) Punto medio de la Clase (x)	Frecuencia (No. De Niños)	•. × •
0 – 49.999	25.00	99	2475
50 – 99.999	75.00	43	3225
100 – 149.999	125.00	85	10625
150 – 199.999	175.00	65	11375
200 – 249.999	225.00	78	17550
250 – 299.999	275.00	15	4125
•		385	49375

Fuente: Los Autores

$$\bullet = \frac{\bullet\dot{.}..\bullet\bullet. \times \bullet.\bullet}{\bullet}$$

$$\bullet = \frac{\bullet\bullet\bullet\bullet\bullet}{\bullet\bullet\bullet} = \bullet\bullet\bullet.\bullet\bullet$$

• = •••.•• • •••• ó ••• ••••• •••••••• •• ••ñ•

2.14.1.2 Propiedades de la Media Aritmética

- Tiene un cálculo sencillo en el cual intervienen cada uno de los datos que se encuentran en la muestra ó población

Se utiliza para comparar frecuencia entre poblaciones, pero se recomienda utilizar este con una medida de dispersión

- Es la interpretación de "punto de equilibrio" ó en otros casos de "centro de masa", ya que puede equilibrar las desviaciones de sus datos con relación a su propio valor.

•∴.(•. • •) / • = •∴. •. / • • •∴. • / • = • • • = •

- El valor de una desviación cuadrática es mínimo cuando;

  •∴.(•.••)• / • • • = •, esto se conoce como Teorema de Kôning, y es muy importante en la varianza

  Su valor se ve influido por cambios de origen y escala quiere decir que:

  Si •.•= ••• + • entonces •• = •• + • , donde ••es la media aritmética de los •.•, para • = 1, . . . , • y • y • números reales.

- Es poco sensible a tener fluctuaciones o variabilidad. Es un parámetro muy útil en la estadística
- Cuando en un conjunto de datos existe mucha dispersión; sucede que entre más heterogéneos sean los datos menor va ser representativos la media de sus componentes
- La media se ve muy afectado por los valores extremos, y los valores que están por encima o son superiores siempre van a tener mayor peso, por

ejemplo si se está realizando un experimento en un laboratorio con cultivo de bacterias y existen 6 muestras; en la muestra numero 1 se produjeron 50 bacterias, y en las muestras 2,3,4,5,6 se produjeron solamente 10 esto quiere decir que el valor de la muestra numero 1 tiene tanto peso como el de las otras 5 muestras, y su media será 16,666 valor que se encuentra por encima de la mayoría de los datos de la población

2.14.2 La media ponderada:

La media pondera es un tipo de media, en la cual se calcula el promedio dándole importancia (mayor – menor) ó *peso* de cada valor con relación al total

La formula de la media ponderada es la siguiente:

••• = • •.•.• • •

Tabla 2.12 Notaciones de la media ponderada

Notaciones
••• = media ponderada
• =Sumatoria de un conjunto de valores
• =peso asignado a cada una de las observaciones
•(• •) = la sumatoria de los productos de los pesos con cada elemento correspondiente
• • = Suma de todos los pesos de las observaciones

Fuente: Estadística y probabilidad de Ciro Martínez B.

Ejemplo 2.8:

En una empresa de Obras Civiles FAPF se han realizado 7 tomas de datos relacionadas con el ACPM suministrado en una motoniveladora, y su costo en las últimas 2 semanas, para esto se suministró los datos en la siguiente tabla:

Tabla 2.13 Costo Vs Cantidad de Galones de Gasolina

Suministro	Costo por Galón	Cantidad de Galones
1	5850	35
2	6500	25
3	4200	75
4	5750	45
5	6200	15
6	5400	54
7	6550	8

Fuente: Los Autores

Se observa que el costo del galón de varia de $4200 a $6550 y la cantidad de 8Gl a 75 Gl, si se desea hallar el costo promedio por galón de ACPM, se aplica media ponderada:

Los valores del costo en los datos son

$$\bullet\bullet = \frac{\bullet \quad \bullet_{\bullet}\bullet_{\bullet}\bullet \quad \bullet \quad \bullet}{\bullet\bullet\bullet}$$

$$\bullet_{\bullet} = 5850;\ \bullet_{\bullet} = 6500;\ \bullet_{\bullet} = 4200;\ \bullet_{\bullet} = 5750;\ \bullet_{\bullet} = 6200;\ \bullet_{\bullet} = 5400;$$
$$\bullet_{\bullet} = 6550$$

Para este ejemplo tomaremos los factores de ponderación o pesos iguales a los galones correspondientes

$$\bullet_{\bullet} = 35;\ \bullet_{\bullet} = 25;\ \bullet_{\bullet} = 75;\ \bullet_{\bullet} = 45;\ \bullet_{\bullet} = 15;\ \bullet_{\bullet} = 54;\ \bullet_{\bullet} = 8$$

Y tenemos que

$$\bullet\bullet = \frac{\bullet(\bullet\bullet\bullet\bullet \times \bullet\bullet) + (\bullet\bullet\bullet\bullet \times \bullet\bullet) + (\bullet\bullet\bullet\bullet \times \bullet\bullet) + (\bullet\bullet\bullet\bullet \times \bullet\bullet) + (\bullet\bullet\bullet\bullet \times \bullet\bullet) + (\bullet\bullet\bullet\bullet \times \bullet\bullet) + (\bullet\bullet\bullet\bullet \times \bullet)}{\bullet\bullet\bullet}$$

••• = ••••••• / ••• = $••••/•••••

Se observa que el costo promedio por galón de ACPM es $5010/•••ó•

Los factores de ponderación en los cálculos son determinados, estos dependen de lo que se esté manejando.

2.14.2.2 Media Ponderada en datos agrupados

Para el cálculo de la media ponderada solamente en datos agrupados se toma el punto medio de cada clase como significativo y la frecuencia de la clase, los pesos son definidos por •, y el denominador es igual a la suma de las frecuencias el cual es el tamaño de muestra •, la formulación es la siguiente,

•• = ••••••• / •

Tabla 2.14 Notación de Media Ponderada en datos agrupados

Notaciones
•• = punto medio de la marca de clase
•• =Frecuencia de la clase
• = • •• = tamaño de la muestra

Fuente: Estadística y probabilidad de Ciro Martínez B.

2.14.3 La media geométrica:

La media geométrica actúa cuando se desea conocer una tasa de cambio, tales como: la serie de crecimiento promedio; en estos casos puntuales no se puede utilizar media aritmética porque proporciona resultados equivocados, entonces se utiliza este tipo de media.

Su formulación es la siguiente:

$$\bullet\ \bullet\bullet\bullet\bullet\ \bullet\bullet\bullet\bullet\ \text{é}\bullet\bullet\bullet\bullet\bullet = \sqrt[\bullet]{\bullet_{\bullet},\bullet_{\bullet},\bullet_{\bullet},\ldots,\bullet_{\bullet}}$$

Tabla 2.15 Notación de la media geométrica

Notaciones
• = numero de valores de •
• . • = Media Geométrica

Fuente: Los Autores

Ejemplo 2.9

Se tiene en un banco un CDT, el cual presenta un crecimiento. Suponiendo que inicialmente son consignados a este como apertura $150, y dejamos que tenga un acumulación de interese (diferentes dependiendo el tiempo) en 3 años.

El crecimiento es igual a:

$$1 + \frac{\bullet\bullet\bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet\bullet\bullet}{150}$$

Este es el factor porque se multiplican los ahorros al inicio del año para obtener el saldo al finalizar este.

Si las tasas de intereses son las siguientes: para el primer año es de 1.09%, para el segundo es de 1.13, y para el tercer año es de 1.16; si este valor es buscado por media aritmética se tendría que $(1.09 + 1.13 + 1.16)/3 = 1.13$, que corresponde a una tasa anual de de 19.5% anual, entonces el depósito solamente crecería:

$\$150 \times 1.195 \times 1.195 \times 1.195 = \$181.64$, lo cual sería un error porque el factor de crecimiento es de $168.945, lo que quiere decir que no se puede desarrollar por media aritmética sino media geométrica, donde seria

$$\bullet.\bullet = \sqrt[\bullet]{1.09 \times 1.13 \times 1.16} = 1.12630$$

• . • = 1.12630 • •••••• ••

Se puede observar que la verdadera tasa de crecimiento promedio es de 12.63% anual, lo cual está un poco lejos de la tasa incorrecta de 19.5% anual

2.14.4 La mediana (• .):

Ésta medida de tendencia central es calculada cuando se organizan los datos de mayor a menor, o viceversa, y se busca el valor medio de estos.

Para calcular la mediana de los datos de una muestra o población dada, se tiene que realizar lo siguiente:

- Se clasifican los datos, sea en orden ascendente o descendente.
- Se siguen cualquiera de los siguientes modos:
    a. Si la cantidad de valores es impar, la mediana se localizara en la mitad de todos los números
    b. Si la cantidad de valores es de tipo par, la mediana se calcula tomando los dos números que están en la mitad y sacándole la media a estas.

Ejemplo 2.10:

En una escuela se desea calcular la mediana de las calorías que tienen los almuerzos de los estudiantes; se toma una muestra de 7 datos con relación a los gramos proporcionados:

150 100 50 230 580 450 300; Se organizan los datos en forma ascendente.

50 100 150 $\bullet\bullet\bullet$ 300 450 580

Como $\bullet = 7$ se toma el valor intermedio, así que la media en las calorías que tienen los almuerzos es 230

Ahora supongamos que deseamos hallar de nuevo la mediana de cuanto es el valor de los almuerzos y tenemos lo siguientes valores:

5000 1000 4500 2700 5800 13000; Se organizan los datos en forma ascendente

$$1000 \quad 2700 \quad \underbrace{\bullet\bullet\bullet\bullet \quad \bullet\bullet\bullet\bullet}_{\text{Dos valores intermedios}} \quad 5800 \quad 13000$$

Como $\bullet = 6$ es par, ya identificados los dos valores intermedios, la mediana seria la media de los números:

$$\bullet_{\bullet} = \frac{4500 + 5000}{2} = \$4750$$

La mediana tiene mejor indicación que la media, porque esta no se deja influenciar por valores extremos.

2.14.4.1 Mediana en Datos Agrupados:

Cuando se tiene un grupo de datos agrupados de tipo par en una distribución de frecuencia y se desea hallar la mediana se debe saber cuál es la sumatoria de las frecuencias, y seria hallar $\frac{(\bullet\bullet\bullet)}{\bullet}$, después se determina el numero de elemento dentro de la clase y la localización de la clase que contiene a la mediana, seguidamente se determina el ancho de cada paso para pasar de una observación a otra en la clase mediana, dividiendo el intervalo de la clase entre el número de elementos contenido en la clase, después se determina el numero de pasos que hay desde el límite inferior de la clase mediana hasta el elemento correspondiente a la mediana, se calcula el valor estimado de la mediana multiplicando el numero de pasos necesarios para llegar a la observación mediana por el ancho que se le da a cada paso y al producto se le suma el límite inferior de la clase.

Cuando el numero de clase es impar simplemente se aplica la siguiente formula

$$\bullet_{.} = \bullet \frac{\frac{(\bullet + 1)}{2} \bullet (\bullet + 1)}{\bullet_{.}} \bullet \bullet + \bullet_{.}$$

Tabla 2.16 Notación de la mediana para datos agrupados

Notaciones
$\bullet_{.}$ =mediana de la muestra
$\bullet$= numero total de elementos de la distribución
$\bullet$= suma de todas las frecuencias de la clase hasta, pero sin incluir la clase de la mediana
$\bullet_{.}$ =suma de todas las frecuencias de la clase mediana
$\bullet$= ancho de intervalo de clase
$\bullet_{.}$ =limite inferior del intervalo de clase de la media

Fuente: Estadística y probabilidad de Ciro Martínez B.

La mediana es más fácil de entender y se calcula con cualquier tipo de datos, incluyendo los datos agrupados con clase de extremos abiertos en distribución de frecuencias, a menos que la mediana se encuentre en esos datos, pero igual posee desventajas como el tipo de procedimiento que utiliza la mediana, además los datos deben ser ordenados antes de llevar a cabo cualquier calculo

2.14.5 La moda

Este es el valor que se encuentra con mayor frecuencia en los datos, ésta es comúnmente utilizada en distribuciones de tipo asimétrico, para eliminar los efectos de los valores extremos.
También se puede dar el caso que se presente dos modas en un conjunto, y este tipo de moda se conoce como Bimodal.

Algunos autores catalogan la mitad del rango como una medida de tendencia central, y es la que constituye el valor que se encuentra en la mitad de una serie de valores y esta se halla realizando una suma del valor máximo y del mínimo y dividiéndolo entre 2.

$$mitad\ del\ rango = \frac{(valor\ maximo + valor\ minimo)}{2}$$

Tabla 2.17 Datos sobre la Moda

PARA TENER EN CUENTA	
	•Cuando una coleccion de datos es aproximadamente simetrica, con una moda, la media, mediana y la mitad del rango tienden a ser iguales •Cuando los datos es de tipo asimetrico, es bueno probar la media y la mediana

Fuente: Los Autores

2.14.5.1 Calculo de la moda en datos agrupados:

Cuando los datos se encuentran en una distribución de frecuencia, es de suponer que la moda se localiza en la clase que contiene la mayor frecuencia. Por consiguiente se utiliza la siguiente formula

$$x_{Mo} = L_{Mo} + \left(\frac{d_1}{d_1 + d_2}\right) A$$

Tabla 2.18 Notación de la moda en datos agrupados

Notaciones
d_1=frecuencia de la clase modal menos la frecuencia de la clase que se encuentra inmediatamente menor a ella
d_2=frecuencia de la clase modal menos la frecuencia de la clase que se encuentra inmediatamente mayor a ella
A= ancho de intervalo de clase
L_{MO}=Limite inferior de la clase modal

Fuente: Estadística y probabilidad de Ciro Martínez B.

La moda se puede utilizar para todo tipo de datos, no afectan los valores extremos, aunque cuando el valor extremo es muy alto o bajo se escoge el valor más frecuente del conjunto de datos como el valor modal, la moda no se utiliza muy a menudo, otra desventaja es cuando existen tres modas en un conjunto es muy difícil la interpretación de los datos

2.15 MEDIDAS DE VARIABILIDAD

Entre las medidas de variación se encuentran las siguientes:

2.15.1 El Rango (*r*):

En un conjunto de datos es la diferencia entre el valor máximo y el valor mínimo, su fórmula es la siguiente:

••••• = ••••• •• •• • •• •••••• • ••••• •• •• • •• •••••ñ•

Es fácil encontrar el rango, pero su utilidad es muy limitada, ya que este solo tiene en cuenta el valor más alto y el valor más pequeño.

2.15.2 Rango Cuartílico (*Q*)

En el momento de ordenar los datos de una población en función al valor de la variable, se dividen en grupos de intervalos iguales, de tal forma que cada uno de ellos tenga el mismo número de observaciones, los cuartiles son valores de la variable de los elementos que ocupen las posiciones divisorias.

Este tipo de rango se clasifica por el numero de partes en que se desee dividir la distribución, los cuales son: Cuartiles son equivalente a cuatro partes iguales; Quintiles que son equivalentes en cinco partes iguales; percentiles o centiles en cien partes iguales

Para obtener los valores de los Cuartiles se debe utilizar:

$$\bullet_{\bullet} = \bullet_{\bullet\bullet\bullet} + \frac{\bullet\frac{\bullet\ \bullet}{4}\bullet\ \bullet\ \bullet_{\bullet\bullet\bullet}}{\bullet_{\bullet}}\bullet_{\bullet}$$

Q_k= Es el valor de la variable cuartil K (K=1,2,3, o 4)

Los quintiles (C_k) se pueden utilizar la siguiente expresión general:

$$\bullet_{\bullet}\ \bullet\ \bullet_{\bullet\bullet\bullet} = \frac{\bullet\frac{\bullet\ \bullet}{\bullet}\bullet\ \bullet\ \bullet_{\bullet\bullet\bullet}}{\bullet_{\bullet}}\bullet_{\bullet}$$

C_k= Es el valor de la variable para el cuartil de orden K de C

Para hallar el valor del percentil

$$\bullet_{\bullet} = \bullet_{\bullet\bullet\bullet} + \frac{\bullet\frac{\bullet\ \bullet}{100}\bullet\ \bullet\ \bullet_{\bullet\bullet\bullet}}{\bullet_{\bullet}}\bullet_{\bullet}$$

P_k= Es el valor de la variable para el percentil K pedido

Ejemplo 2.11:

Hallar los Cuártiles de la siguiente distribución:

Tabla 2.19 Datos de Distribución

x_i	18	19	20	21	22
n_i	10	30	35	15	10
N_i	10	25	50	75	100

Fuente: Los Autores

Se divide la distribución en cuatro partes iguales: Q_1, Q_2, Q_3 y Q_4, donde hay realmente tres Cuartiles significativos. Pensemos que $Q_{1,}$ será el primer cuarto de observación y así sucesivamente,

Los Cuártiles serán los valores de la variable a la cual corresponde la frecuencia absoluta acumulada $\bullet/4, 2\bullet/4\ \bullet 3\bullet/4$ y la distribución es $\bullet = 200$, tendremos que $\bullet/4 = 50$; $2\bullet/4 = 100$; $3\bullet/4 = 150$

Los valores de la variable que corresponde a cada uno de los cuartiles es:

$$\bullet 1 = 19;\ \ \bullet 2 = 20;\ \ \bullet 3 = \frac{20+21}{2} = 20{,}5$$

2.15.3 Rango Interfractil

Es una medida de dispersión entre dos fractiles de una distribución de frecuencia, diferencia entre los valores de los dos fractiles en una distribución de frecuencia, que proporciona datos que caen en un fractil o debajo de este; un ejemplo clave es la mediana el cual tiene un fractil de 0.5 por que la mitad del conjunto de datos en ella es menor o igual a este valor.

Los fractiles tienen nombres especiales, dependiendo del número de partes en las que se dividen los datos, los fractiles que se dividen en 10 se llaman deciles, percentiles los que se dividen en 100

2.15.4 Rango Intercuartil

El cuartil son los valores más altos de cada una de estas cuatro partes, y el rango Intercuartil es la diferencia entre los valores del primero y tercer cuartil

$$\bullet\bullet\bullet\bullet\bullet \qquad = \bullet_{\bullet}\ \bullet\ \bullet_{\bullet}$$

El rango además mide aproximadamente que tan lejos esta de la mediana, y se debe ir en cualquiera de las dos direcciones antes de recorrer una mitad en los valores de los rangos del conjunto de datos. Para cálculo de este rango se divide en cuatro partes, en otras palabras, el 25% de los datos de la distribución

2.15.5 La Desviación Estándar

Es la medida de variación de los valores con respecto a la media. El valor de la desviación estándar es un valor positivo, y la única posibilidad de que sea igual a cero es que los valores de los datos sean iguales, porque no presentar una variación. Para calcular la desviación estándar se utiliza la siguiente ecuación:

$$\bullet = \bullet\overline{\frac{\bullet(\bullet \bullet \bullet)^{\bullet}}{\bullet \bullet 1}}$$

La fórmula de cálculo de la desviación se obtiene de la siguiente manera:

$$\bullet^{\bullet} = \frac{\bullet(\bullet \bullet \bullet)^{\bullet}}{\bullet \bullet 1} = \frac{\bullet(\bullet^{\bullet} + \bullet^{\bullet} \bullet 2\bullet\bullet)}{\bullet \bullet 1} = \frac{\bullet \bullet^{\bullet} + \bullet\bullet^{\bullet} \bullet 2\bullet \bullet \bullet}{\bullet \bullet 1}$$

Y puesto que $\bullet = (\frac{\bullet}{\bullet}) \bullet \bullet$, la cual se reduce a:

$$\bullet = \bullet\overline{\frac{\bullet(\bullet \bullet \bullet)^{\bullet}}{\bullet \bullet 1}}$$

Para hallar la desviación estándar de la muestra se utiliza la siguiente expresión:

$$\bullet = \bullet\overline{\frac{\bullet \bullet(\bullet^{\bullet})\bullet(\bullet \bullet)^{\bullet}}{\bullet(\bullet \bullet 1)}}$$

Las unidades de la desviación estándar son las mismas unidades en las que se encuentran los datos observados.

También existe una desviación estándar para la población, su fórmula es la siguiente:

$$\bullet = \bullet\overline{\frac{\bullet(\bullet \bullet \bullet)^{\bullet}}{\bullet}}$$

2.15.6 Varianza de una Muestra y una población.

La varianza es una medida de variabilidad, la cual es igual al cuadrado de la desviación estándar, esta presenta una gran desventaja: las unidades de la varianza son diferentes a las unidades del conjunto original de datos, esto se debe a que las unidades de la desviación estándar también son elevadas al cuadrado en el momento de buscar la varianza.

Existen dos ámbitos en los cuales se puede buscar la varianza, los cuales son:

Varianza muestral: es el cuadrado de la desviación estándar de la muestra s^2.

Varianza poblacional: es el cuadrado de la desviación estándar poblacional $\bullet^{2}$

En todos los casos la varianza muestral utiliza el divisor • • 1, es decir el tamaño de la muestra menos1; por otra parte la varianza poblacional implementa como divisor el tamañ006F de la población •, si se conociera el valor exacto de la media poblacional •, de esta manera se encuentra la varianza muestral elevando al cuadrado la desviación promedio de las observaciones muéstrales alrededor de •.

2.15.6.1 Propiedades de la Varianza

Es de gran importancia conocer las propiedades de la varianza, ya que en algunos casos simplifica las operaciones, además de servir como base de algún proceso matemático. Las propiedades son las siguientes:

- Ø La varianza siempre debe ser positiva. • 0.
- Ø La varianza de una variable mas una constante es igual a la varianza de la variable.

$$\bullet_{[\bullet\bullet\bullet]} = \bullet_{[\bullet]} + \bullet_{[\bullet]} = \bullet\colon$$

- Ø La varianza de una constante por una variable es igual al producto de la constante elevada al cuadrado por la varianza de la variable.

$$\bullet_{[\bullet\bullet\bullet]} = \bullet^{\cdot}\bullet\colon$$

Es una propiedad que facilita la conversión de la varianza, sin la necesidad de volver a retomar todos los datos iníciales;

Ejemplo: En un estudio la media es 10 horas y su desviación es de 4,2 horas, se desean convertir esos resultados en minutos, para mayor rapidez en esto se utiliza esta propiedad de la varianza

• $_{(km)}$ = • • = 60(10) = 600 minutos, ahora se calcula la varianza

• $_{[x.\ k]=\ K}$ $^{2}S^{2}$ = 602(4,2) = 15.120

Ø La varianza de dos o más sub-muestras se obtiene mediante la siguiente aplicación:

$$\bullet^{\bullet} = \frac{\bullet \bullet \bullet + \bullet \bullet \bullet}{\bullet} + \frac{\bullet(\bullet \bullet \bullet)^{\bullet}\bullet + \bullet(\bullet \bullet \bullet)^{\bullet}\bullet}{\bullet}, \quad \bullet = \bullet + \bullet \bullet 2$$

Hay que tener en cuenta que: $\bullet\bullet = \frac{\bullet\bullet\bullet + \bullet\bullet\bullet}{\bullet}$

Supongamos que los datos para la variable discreta que se utilizan corresponde a la primera sub-muestra con la media • = 2,82 y de varianza • = 2,19 , mientras que la segunda sub-muestra es • = 51,8 y • = 59,81, con esta información se calcula la media y la varianza, para un conjunto de 101 observaciones cada una. Siendo:

$$\bullet\bullet = \frac{2{,}82(100) + 51{,}8(100)}{200} = 27{,}31\ \bullet$$

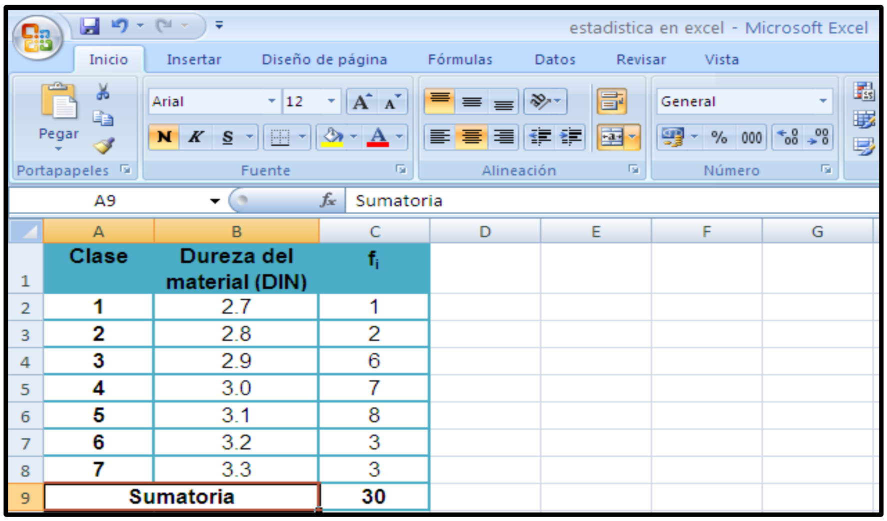
$$\bullet = 101 + 101 \bullet 2 = 200$$

$$\bullet^{\bullet} = \frac{2{,}19(100) + 59{,}81(100)}{200} + \frac{(2{,}82 \bullet 27{,}31)^{\bullet}(100) + (51{,}8 + 27{,}31)^{\bullet}(100)}{200}$$

$$\bullet^{\bullet} = 31 + 599{,}7601 = 630{,}7631$$

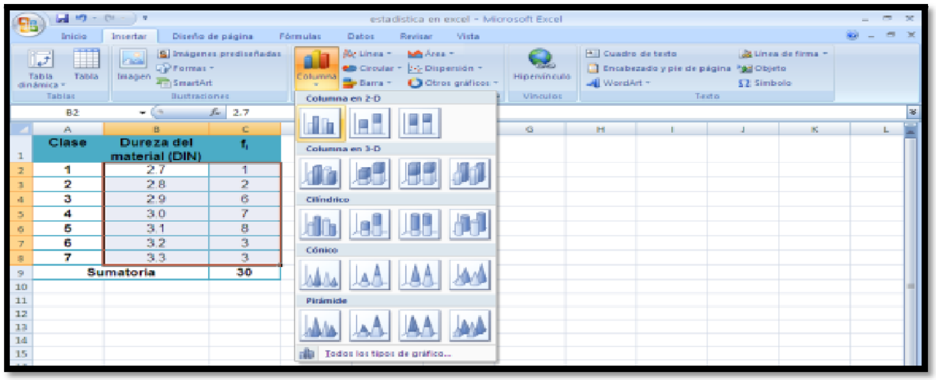
2.16 CONSTRUCCIÓN DE DIAGRAMAS DE BARRAS, POLÍGONOS DE FRECUENCIA Y DIAGRAMA CIRCULAR A TRAVÉS DE EXCEL

1. Inicialmente, se debe ingresar los datos a la hoja de cálculo, para realizar el ejercicio se utilizarán los datos del ejemplo 2.4

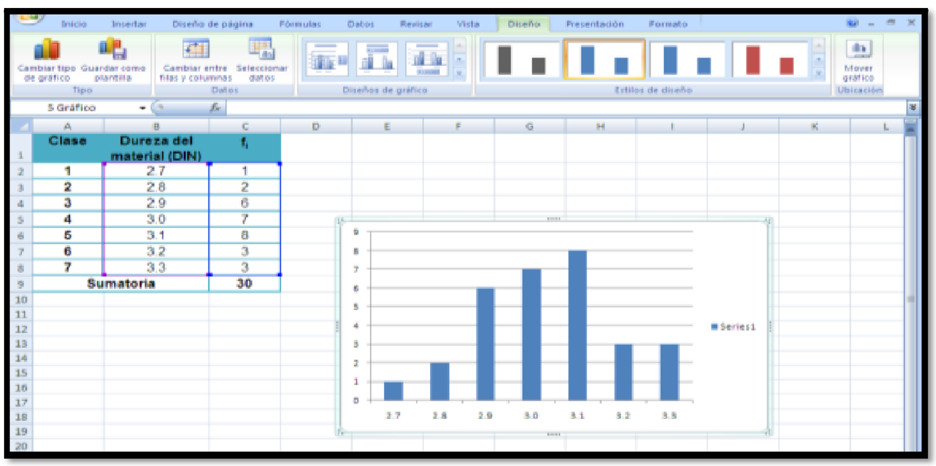


Clase	Dureza del material (DIN)	f_i
1	2.7	1
2	2.8	2
3	2.9	6
4	3.0	7
5	3.1	8
6	3.2	3
7	3.3	3
Sumatoria		30

2. En la barra de herramientas, se escoge la opción insertar se escoge el tipo de gráfico a realizar. Previamente, se debe seleccionar los datos.
   Para realizar el diagrama de barra, se escoge la opción columna y la forma del grafico de barras que se desee. El ejercicio se realizará un gráfico en 2 D. Es decir, en dos dimensiones.



3. Dar clic en el icono que representa la forma del grafico de barras a realizar, automáticamente se visualiza la grafica en la hoja de cálculo.



4. Si se desea realizar cambios al grafico de barras, se da clic sobre la imagen para acceder a la opción herramientas de gráficos, el cual ofrece la siguientes funciones:
   - o Cambiar el diseño y estilo (colores y forma) del gráfico
   - o Modificar la presentación del gráfico, la cual incluye:
     - Adicionar títulos de ejes y del gráfico
     - Visualizar la tabla de datos, leyenda y etiqueta de datos

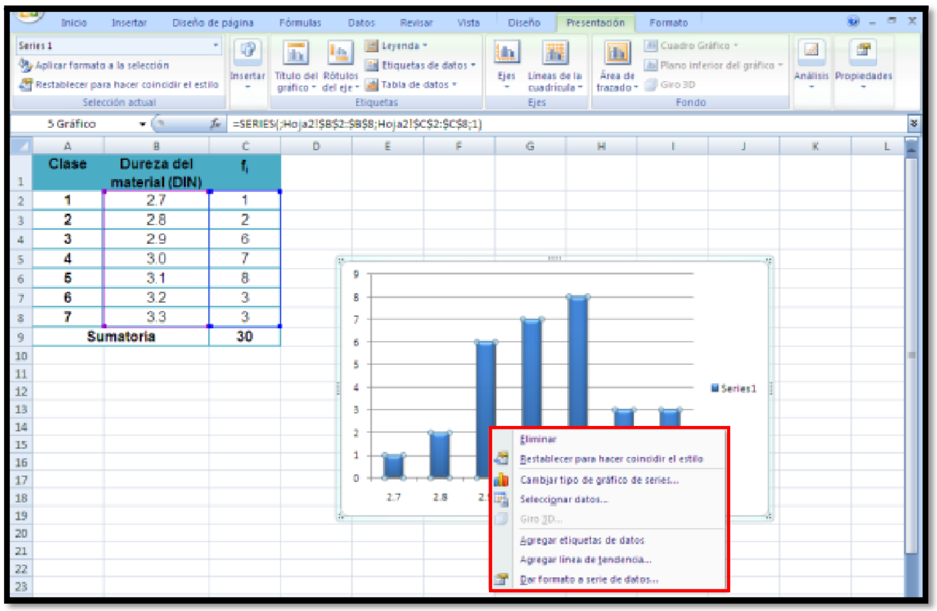
- Cambiar el formato y diseño de cada eje, por medio de la opción eje
- Visualizar área de trazado del gráfico
- Realizar un análisis del grafico
- Verificar las propiedades del gráfico

o Cambiar el formato al gráfico:

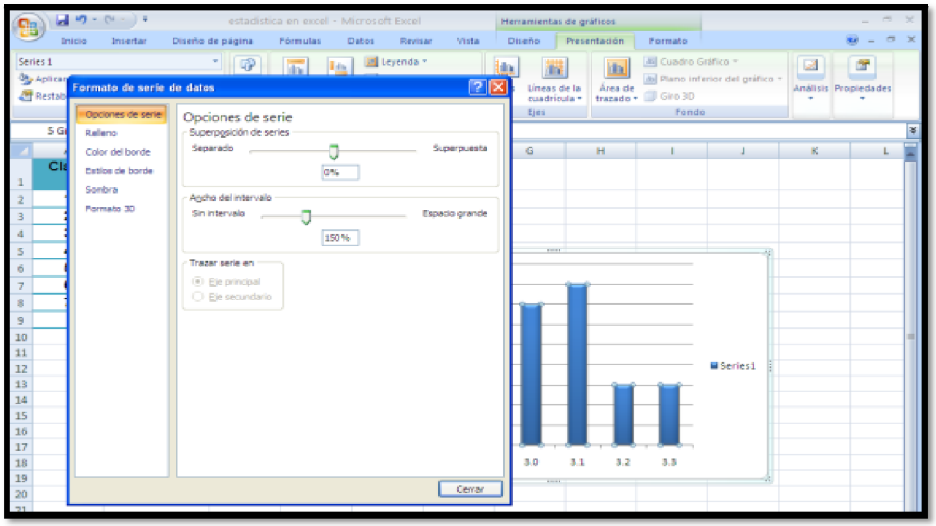
  § Colocar contorno al área del gráfico

  § Modificar los colores de relleno del gráfico

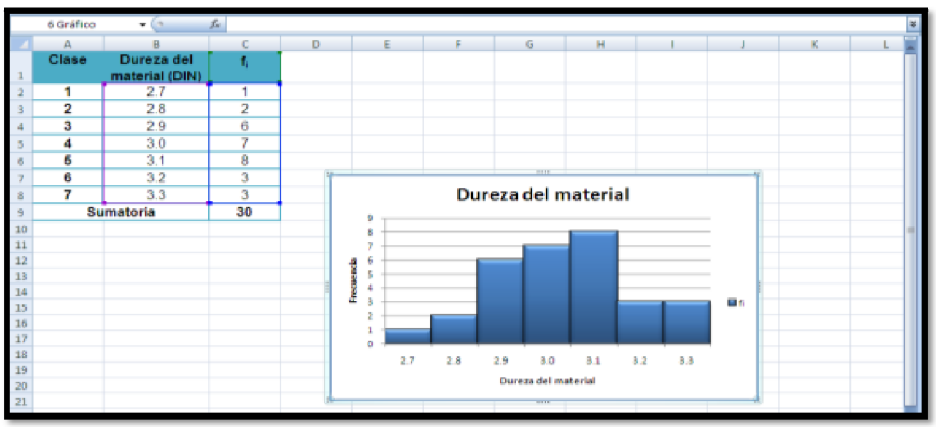
Otros cambios se pueden realizar al dar clic sobre las barras del gráfico para seleccionarlas. Luego, se debe dar clic derecho y se despliega el cuadro con las siguientes opciones:



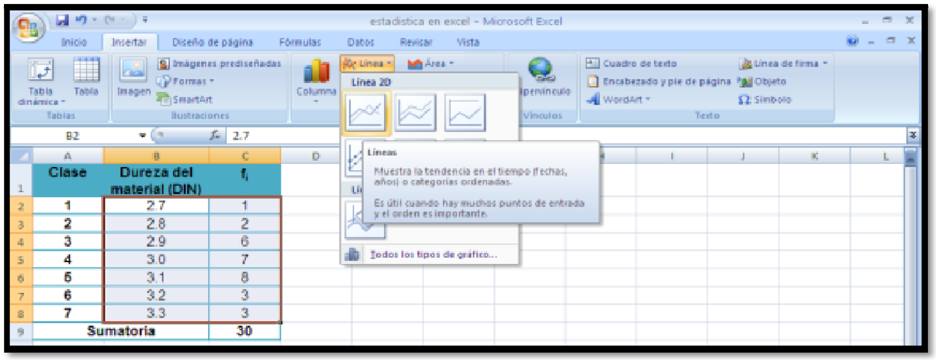
De las opciones se debe escoger Dar formato a serie de datos…

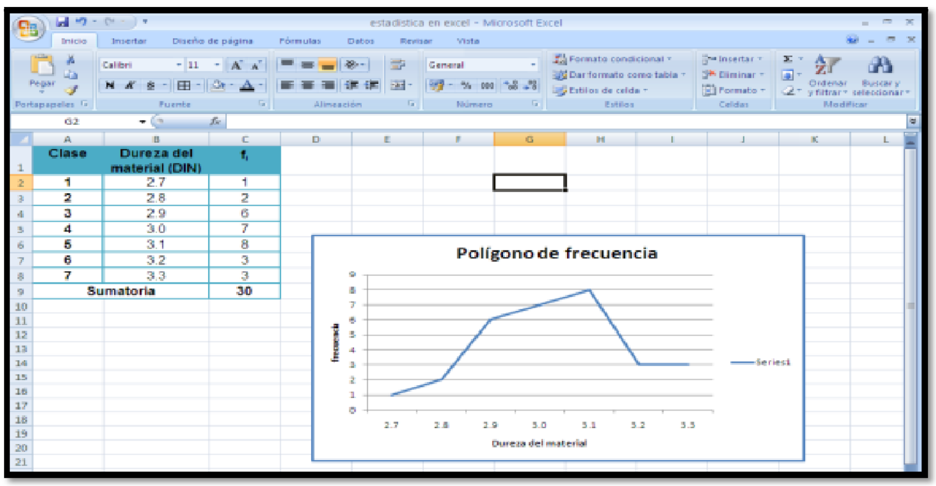


En opciones de serie, se disminuye el ancho del intervalo de las barras y el resto de opciones sirven para modificar el formato de las barras.



Para realizar el Gráfico de polígono de frecuencia se escoge la opción línea y se realiza una línea en 2D.



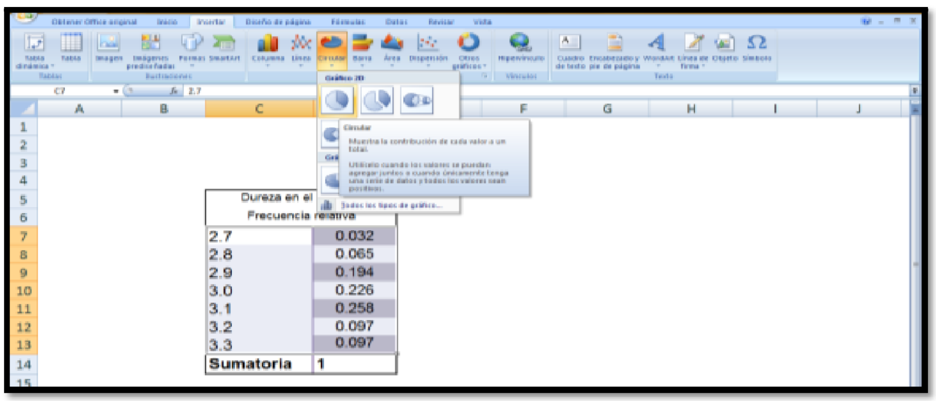


Para realizar el Gráfico de pastel se debe realizar los siguientes pasos:

1. Insertar los datos en la hoja de cálculo.

Dureza en el material Frecuencia relativa	
2.7	0.032
2.8	0.065
2.9	0.194
3.0	0.226
3.1	0.258
3.2	0.097
3.3	0.097
Sumatoria	**1**

2. Se busca en la barra de Insertar el tipo de Grafica que se desea, para realizar el diagrama circular se escoge la opción circular y la forma del gráfico que se desee.



3. Después se inserta la grafica deseada

APLICACIÓN EXCEL

2.17 ANÁLISIS ESTADÍSTICO DE DATOS NO AGRUPADOS A TRAVÉS DE EXCEL

Para realizar el análisis de datos no agrupados a través e Excel se considerará el siguiente ejemplo:

En un estudio realizando con el fin de conocer el peso corporal de un grupo de jóvenes estudiantes de secundaria comprendidas entre las edades de 14 y 17 años, se obtuvieron los siguientes datos:

Tabla 2.20. Datos de Peso en Kilogramos

DATOS (peso en kg)									
57	48	51	43	50	51	45	49	46	52
49	57	47	55	52	46	47	50	52	46
45	46	57	53	45	45	48	51	46	49
47	53	52	48	46	45	47	45	50	46

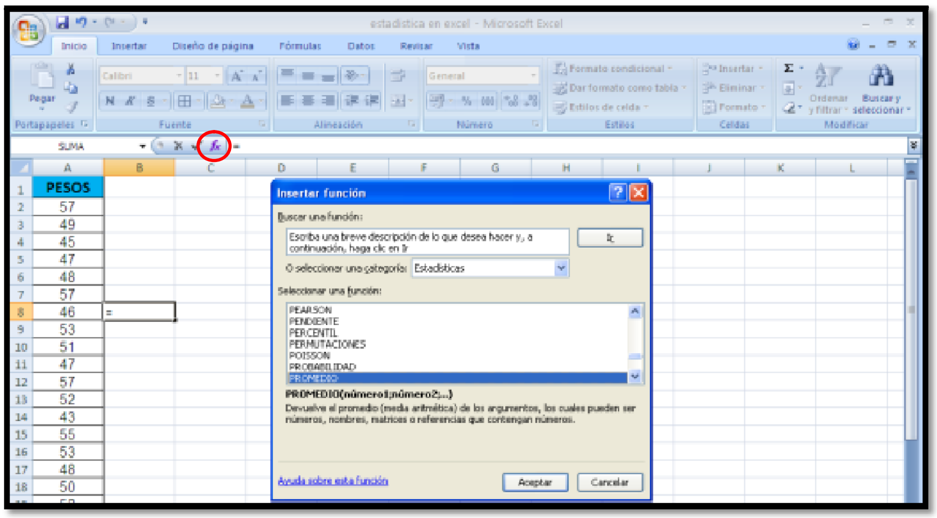
Fuente: Los Autores

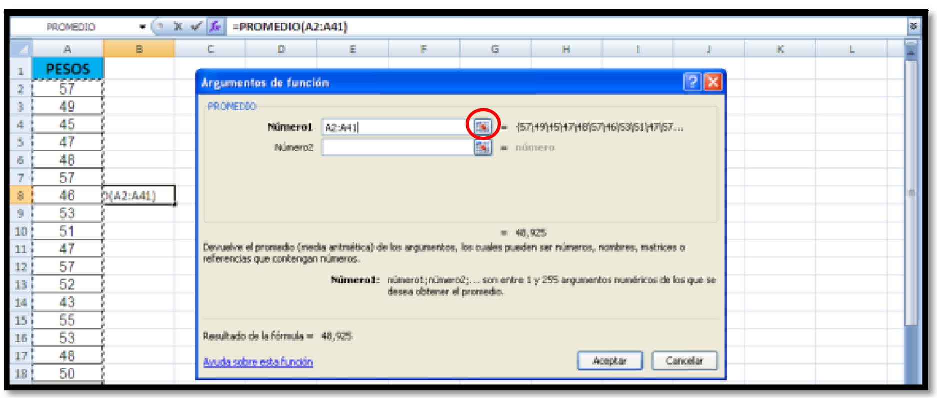
PASOS:

1. Ingresar los datos en una columna.
2. Para hallar la media de los datos, se debe ubicar sobre una celda vacía donde se desee mostrar el resultado y se da clic en el icono f(x) o función.
   Se selecciona la categoría estadística o se puede realizar la búsqueda de la función a través de Buscar una función.

   Para hallar la media de los datos, Excel dispone de una función estadística llamada promedio.

Cuando se selecciona la función promedio y se da clic en aceptar, aparece un cuadro llamado Argumentos de función, en el cual se deben ingresar el rango de valores de la variable peso.

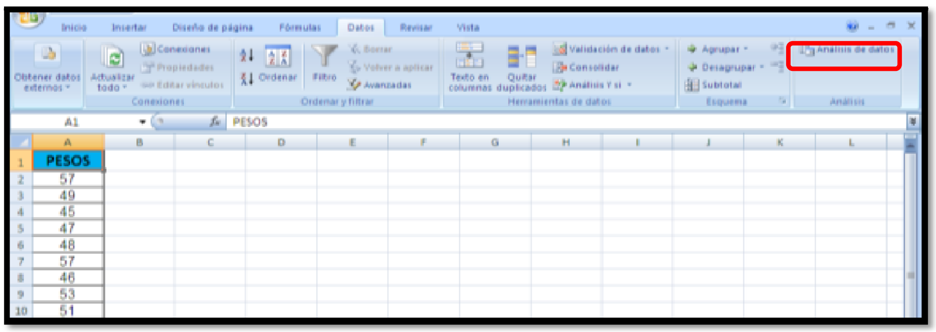




De esta misma manera se obtiene las diferentes medidas de tendencia central y dispersión de los datos.

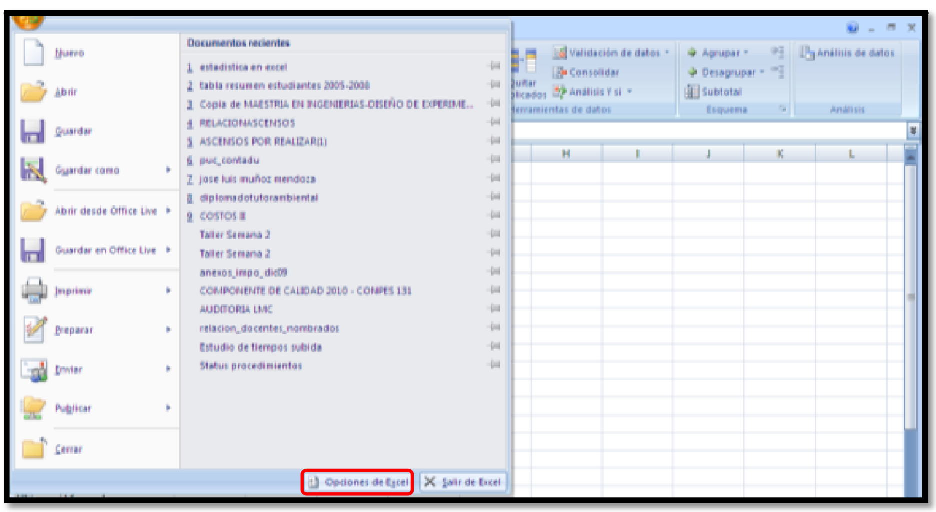
También se puede realizar el análisis estadístico a través de la función complementaria de Excel: análisis de datos.

En la barra de herramientas datos, se verifica que se tenga instalado la función análisis de datos.

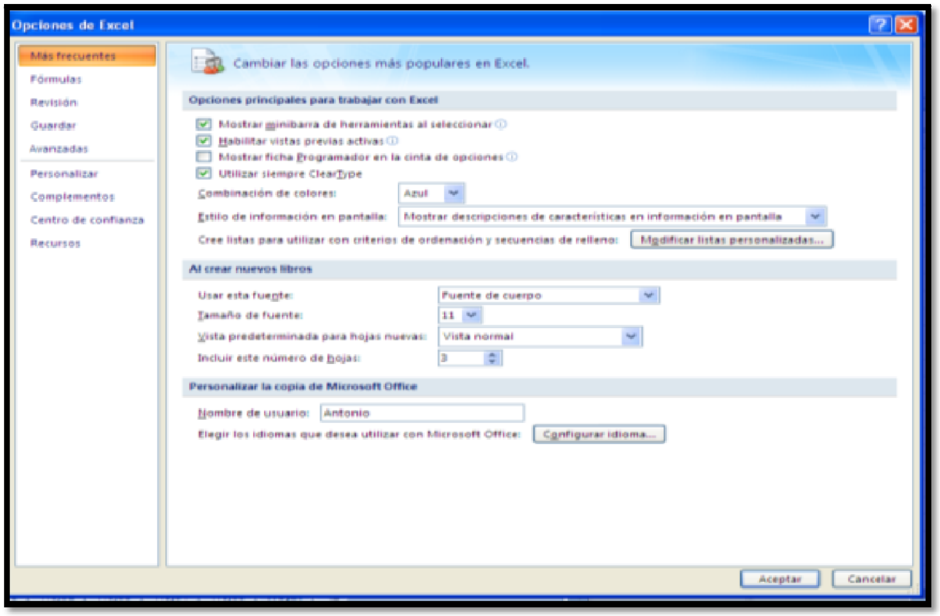


En el caso que no se encuentre instalada, se procede a instalarla, para esto:

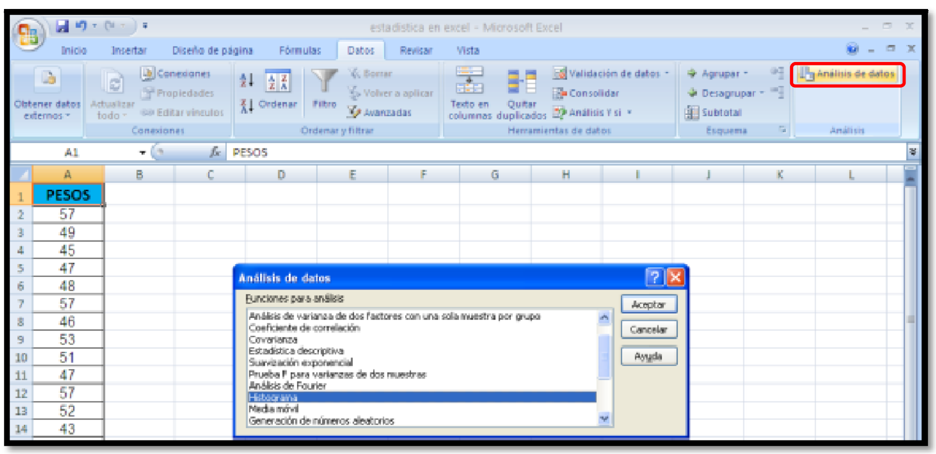
1. Ingresar a opciones de Excel.



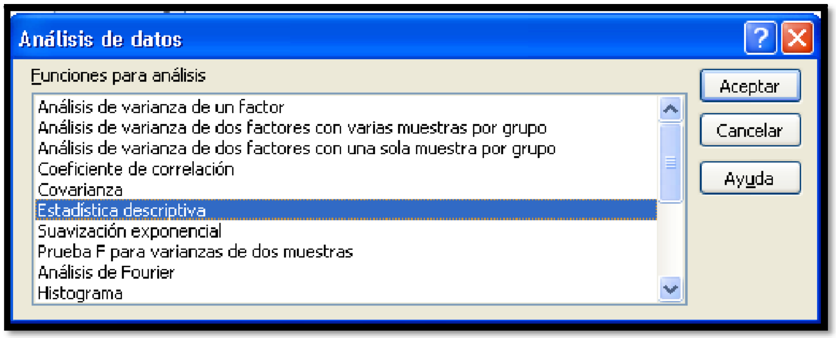
2. En complementos, se escoge la función "*herramientas para análisis*", y posteriormente se realiza la instalación del mismo.



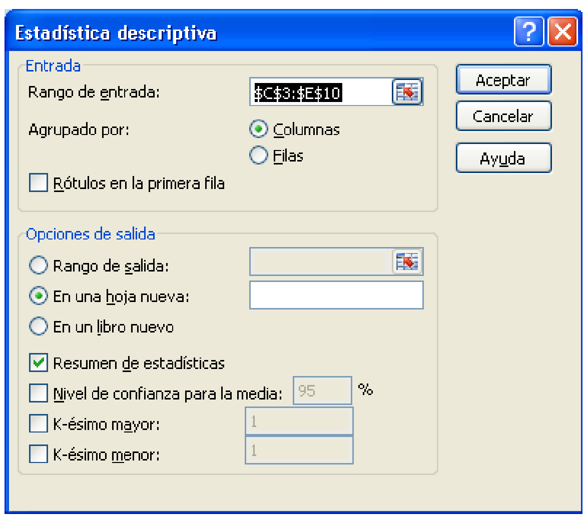
3. Una vez instalada, se da clic sobre el icono análisis de datos y en el cuadro análisis de datos, se escoge la función para análisis.



4. Para realizar un análisis descriptivo de los datos, se escoge la función estadística descriptiva y se da clic en aceptar.



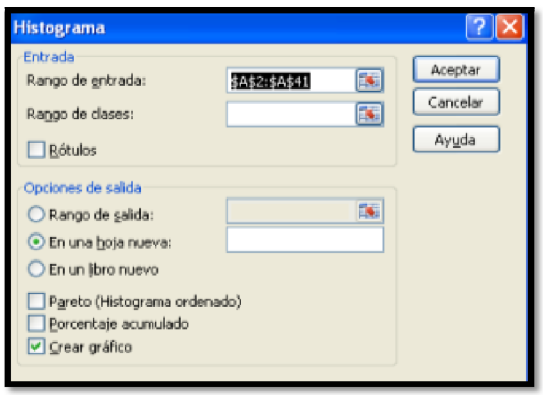
5. Seguidamente se introduce el rango de entrada que corresponde a los datos, y se determina como se encuentra agrupado, teniendo en cuenta el ejemplo que se realiza, los datos fueron ingresados en una columna. Además se debe escoger la opción de salida de los datos que puede ser en la misma hoja de cálculo (Rango de salida), en una hoja nueva o en un libro nuevo.



6. Finalmente, se muestra el resumen estadístico de la siguiente manera:

PESOS	
Media	48,925
Error típico	0,579995027
Mediana	48
Moda	46
Desviación estándar	3,668210631
Varianza de la muestra	13,45576923
Curtosis	-0,22046322
Coeficiente de asimetría	0,715844862
Rango	14
Mínimo	43
Máximo	57
Suma	1957
Cuenta	40

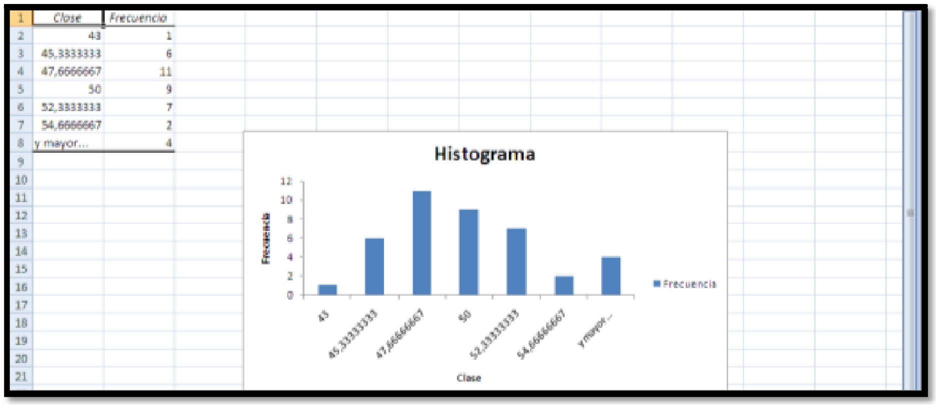
Para realizar el histograma de frecuencia, se escoge la opción histograma del cuadro análisis de datos y se da clic en aceptar.



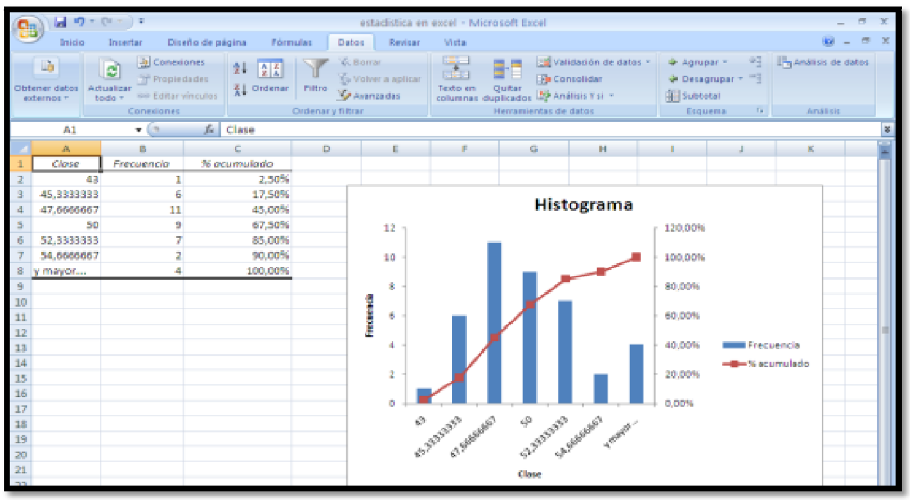
Posteriormente, se escoge todos los datos o valores de la variable y se escoge la opción “crear grafico” en una hoja nueva.

Además ofrece la opción para realizar el histograma con la frecuencia relativa a través de “Porcentaje acumulado”.

Se obtienen los siguientes histogramas:



Clase
Frecuencia
43
1
45,3333333
6
47,6666667
11
50
9
52,3333333
7
54,6666667
2
y mayor...
4
Histograma
Frecuencia
Clase



estadistica en excel - Microsoft Excel
Inicio
Insertar
Diseño de página
Fórmulas
Datos
Revisar
Vista
Clase
% acumulado
Frecuencia
2,50%
17,50%
45,00%
67,50%
85,00%
90,00%
100,00%
Histograma
% acumulado

APLICACIÓN STATGRAPHICS CENTURION

2.18ESTADÍSTICA DESCRIPTIVA EN STATGRAPHICS CENTURIÓN. (Versión XVI)

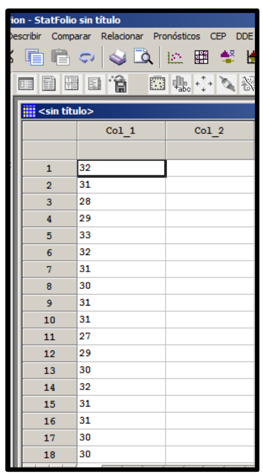
Para ilustrar la forma como se utiliza el software Statgraphics centurión para realizar un análisis estadístico descriptivo, se realizará el siguiente ejemplo:

Durante un turno en una empresa de acero se registraron los siguientes datos de la tenacidad en un material

32,31,28,29,33,32,31,30,31,31,27,28,29,30,32,31,31,30,30,29,29,30,30,31,
30, 31, 34, 33,33, 29,29.

Las unidades de tenacidad son pulg-libra/pulg^3

1. Se introducen los datos a Statgraphics centurión en una columna.



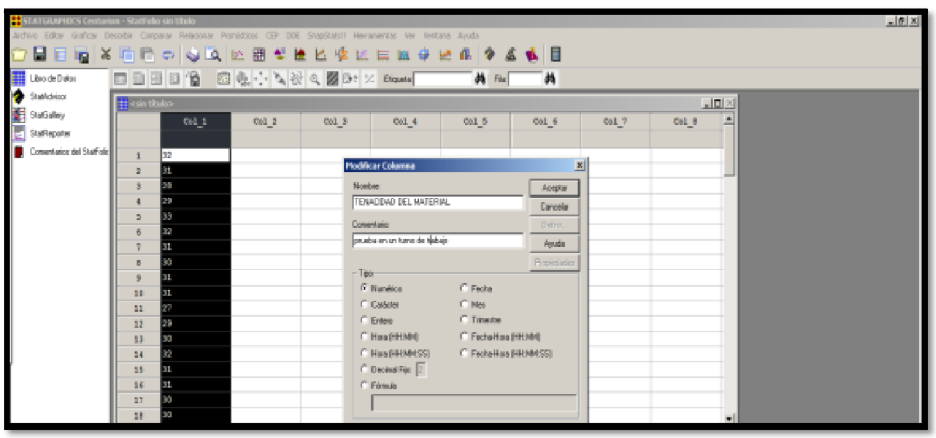
2. Luego se modifica el formato de la columna, para esto se da clic sobre la cabecera de la columna para seleccionarla y luego de da clic derecho y se despliega un cuadro de opciones en la cual se debe escoger modificar columna.

La opción de modificar columna se puede realizar antes de introducir o después de introducir los datos.

3. Después, se despliega el cuadro Modificar columna tal como se muestra en la Figura 3 y se procede a introducir los datos.

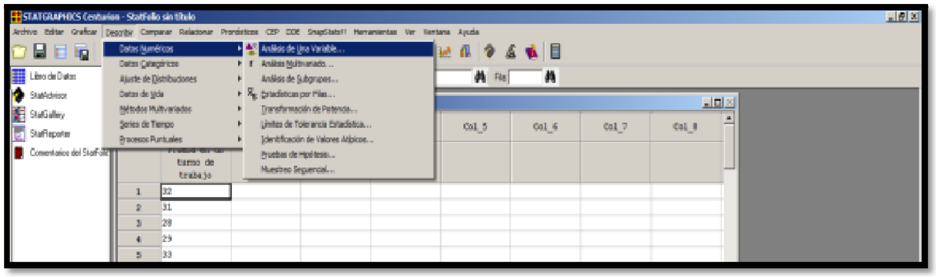
Es importante tener en cuenta que para el tipo de variable numérica, la denotación decimal de los datos es a través de la coma (,), por ejemplo: 2,0-1,2.

Para las variables que representan carácter, los decimales pueden separarse por medio de un punto (.) o una coma (,), por ejemplo: 5,3- 3.2.

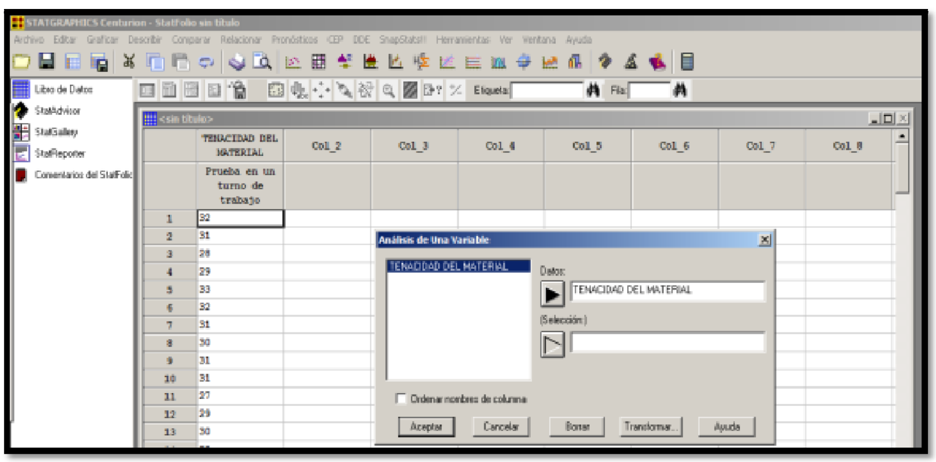


4. Una vez ingresados los datos, se procede a realizar el análisis estadístico de la variable tenacidad del material.

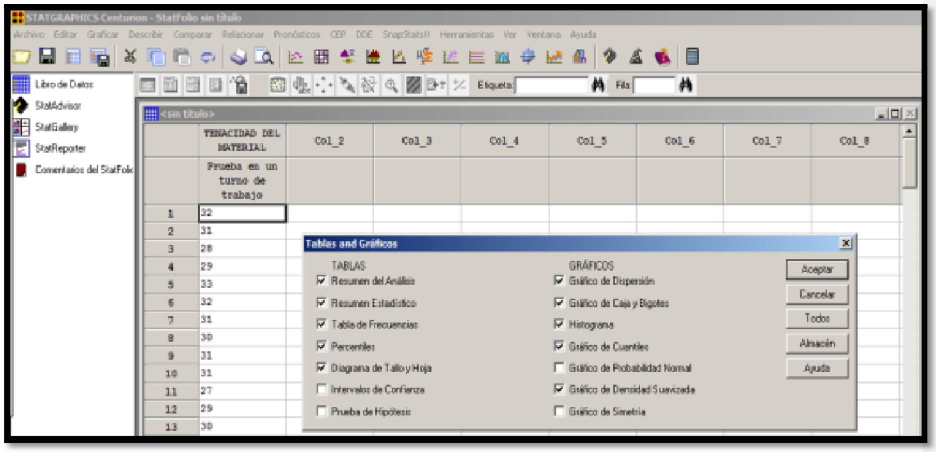
Para realizar una descripciòn estadística detallada del comportamiento de la variable, en la opción Describir, se da clic en la opción Datos Numéricos, finalmente clic en Análisis de una variable.



5. Automáticamente se despliega el cuadro Análisis de una Variable, en el cual se referenciarán los datos a analizar, se da clic sobre el nombre de la variable TENACIDAD DEL MATERIAL, posteriormente se da clic en el icono ▶ del campo datos y se clic en Aceptar.

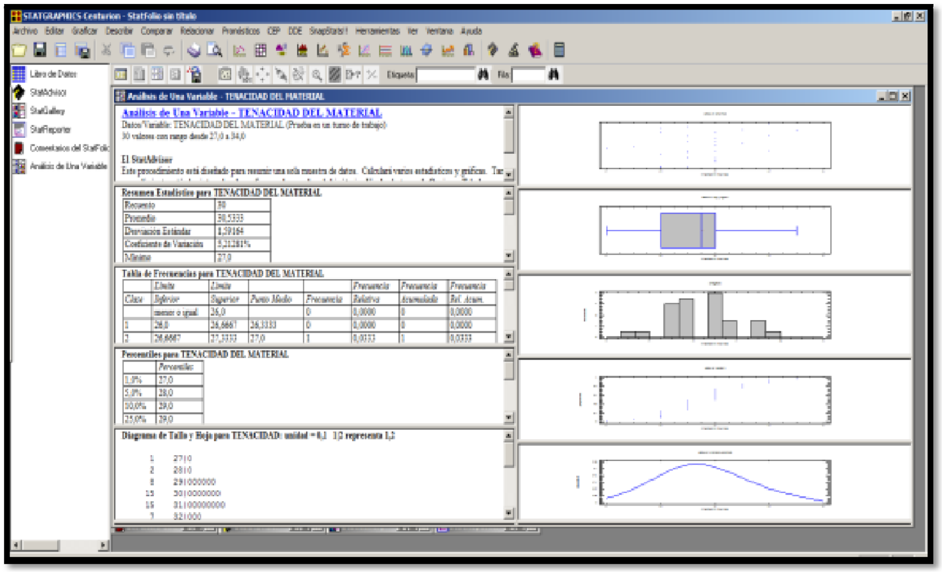


6. En el cuadro Tablas y Gráficos, se escogen las tablas y gráficos que se deseen visualizar y se da clic en aceptar.

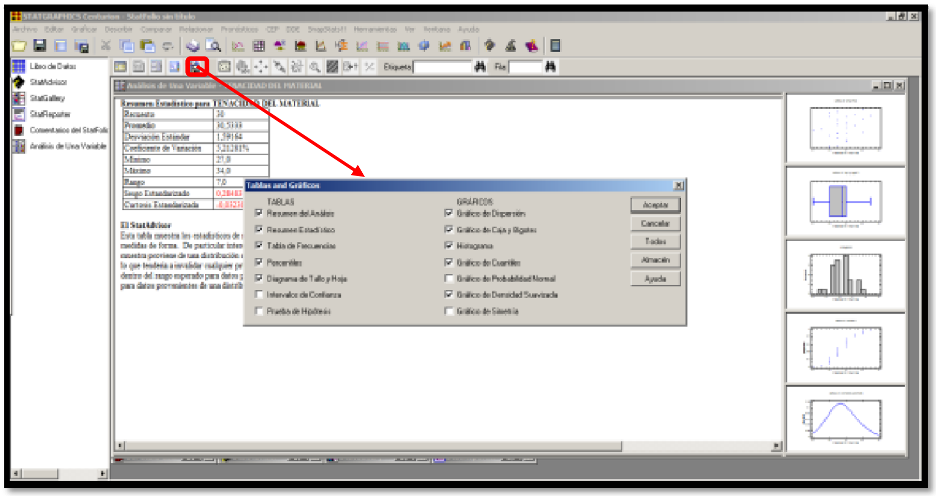


7. Cada recuadro del cuadro Análisis de Una Variable- TENACIDAD DEL MATERIAL muestra la descripción estadística de la variable.

   Si se desea observar de manera detallada, se da doble clic sobre el recuadro y se despliega el recuadro, de igual manera se da doble clic y se minimiza el recuadro.

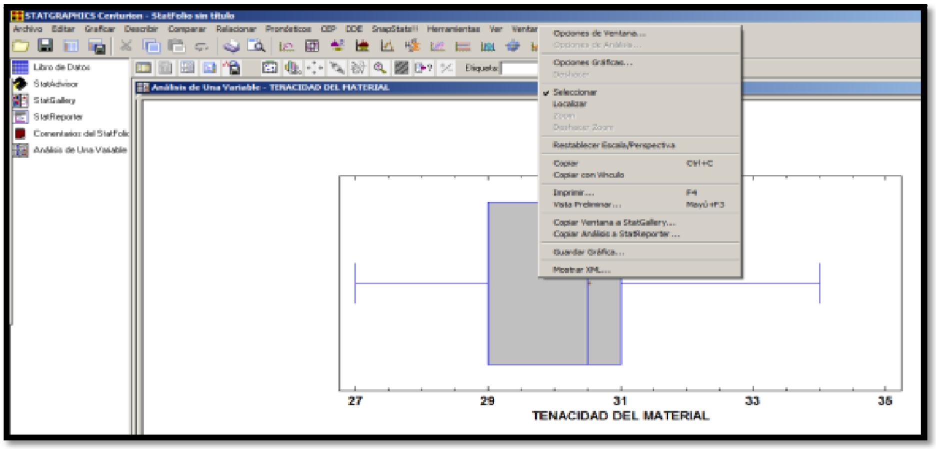


8. Para agregar o desagregar gráficos y tablas se da clic en el ícono 

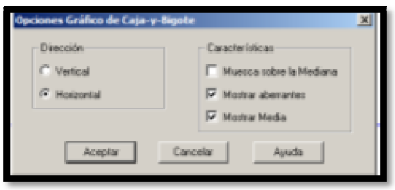


En el recuadro de Resumen Estadístico, se muestra una tabla donde se detalla las medidas de tendencia central y medidas de dispersión de la variable, seguido hay un texto donde se describe de manera general el comportamiento de estas medidas en la variable Tenacidad del material.

Para las opciones de gráficas, tomando el Gráfico de cajas y Bigotes, ubicándose sobre la grafica, se da un clic derecho y aparecen diferentes opciones.



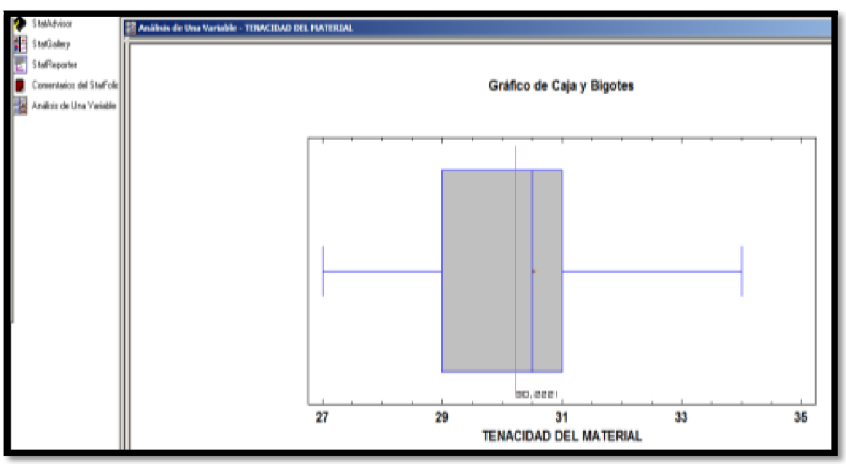
- En opciones de ventana: Se puede cambiar la dirección del gráfico y se modifican las características



.

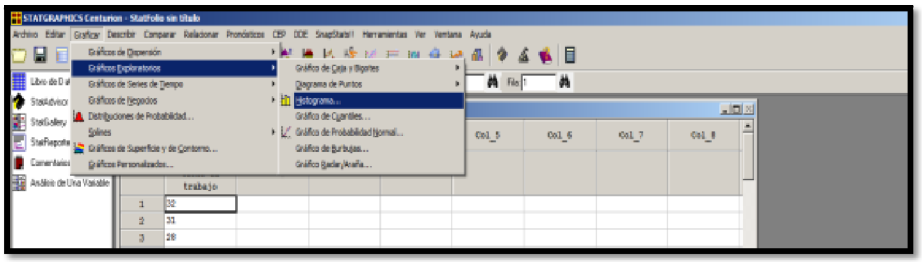
- En Opciones Gráficas: Sirve para modificar el aspecto de la gráfica, relacionados con el diseño, rellenos, título, etc.

- Seleccionar: Como su nombre lo indica, sirve para seleccionar la graficas o parte de ella.
- Localizar: Se activa una línea que tiene la función de localizar un punto en la gráfica.

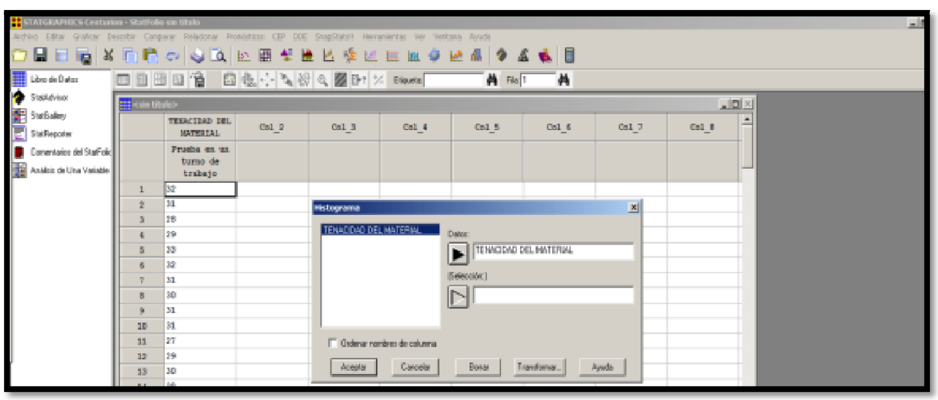


De manera general se pueden utilizar las funciones anteriormente descritas para dar formato a la gráfica.

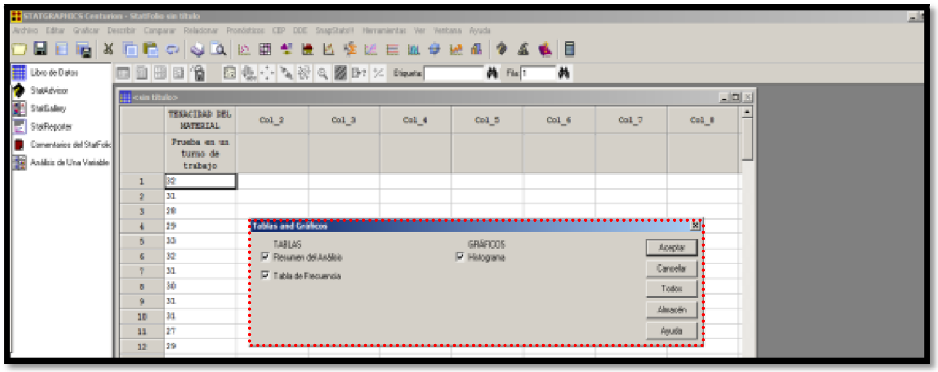
Otra manera de realizar los gráficos es a través del menú Graficar, la cual ofrece diversos tipos de gráficas, para realizar el histograma de frecuencia se utiliza la opción de Gráficos Exploratorios.



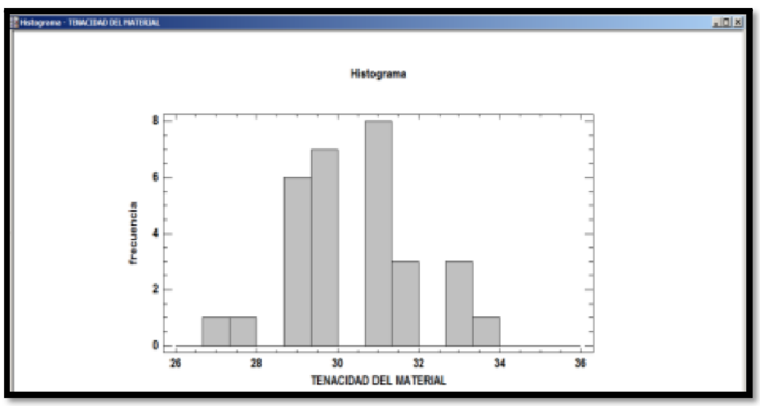
Automáticamente se despliega el cuadro Histograma, en el cual se referenciarán los datos a graficar, se da clic sobre en el nombre de la variable TENACIDAD DEL MATERIAL, posteriormente se da clic en el icono ▶ del campo datos y se clic en Aceptar.



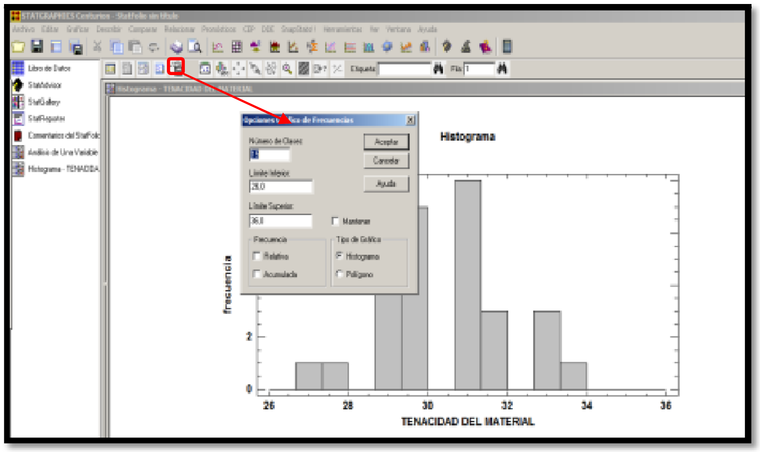
En el cuadro de tablas y gráficos se escoge las tablas y gráficos que se desee visualizar.



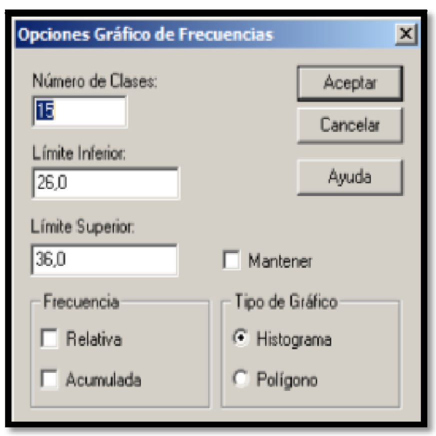
Finalmente se obtiene el histograma de frecuencia.



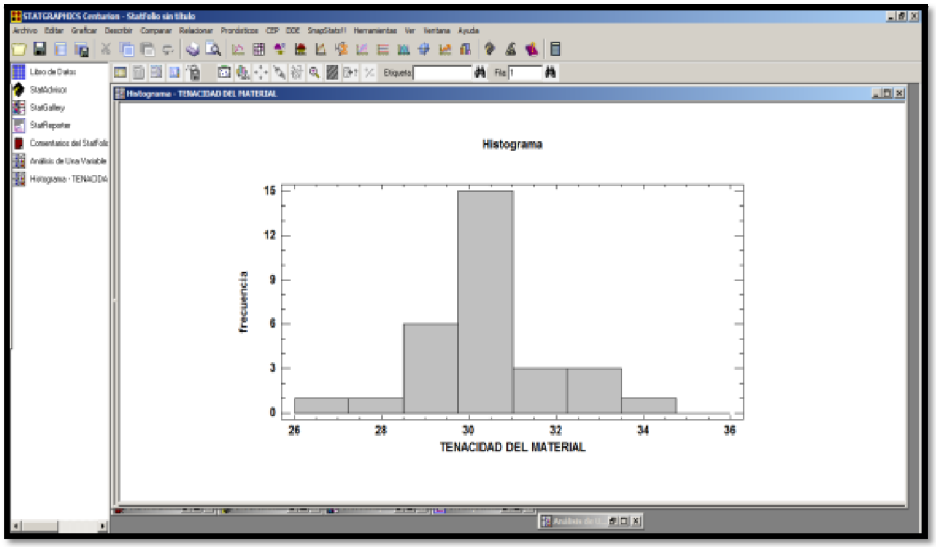
Para modificar las opciones de gráfico, se da clic en el ícono 



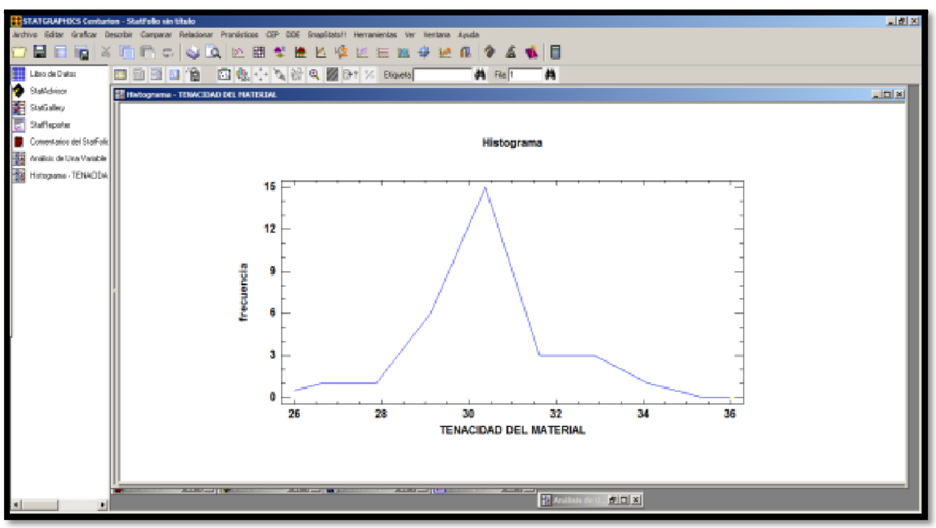
En opciones Gráfico de Frecuencias, se puede modificar el Número de Clases y escoger el gráfico de frecuencia (Relativa, Acumulada) y el tipo de Gráfico (Histograma o polígono).



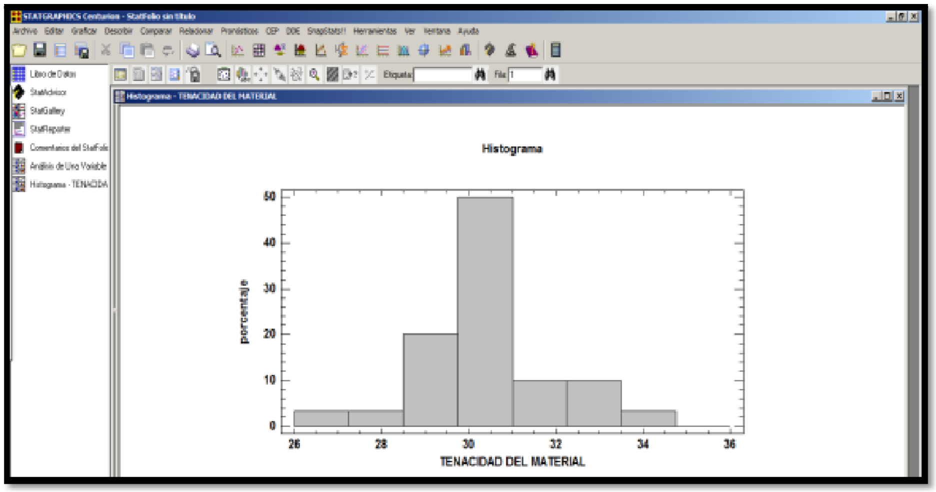
Con un número de clases igual a 8, se obtiene le siguiente histograma:



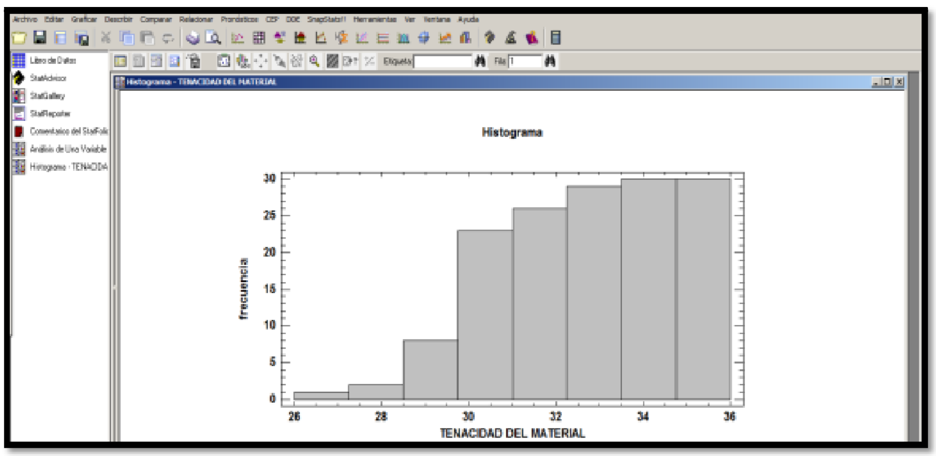
El polígono de frecuencia es el siguiente:



El grafico de frecuencia relativa:



El gráfico de frecuencia acumulada:



De manera general, en Statgraphics centurión se pueden realizar diferentes tipos de gráficos.

EJERCICIOS PROPUESTOS
CAPÍTULO N° 2.

Pregustas propuestas.

Preguntas Tipo I

Pregunta de Selección Múltiple con única respuesta

1. Una población infinita es cuando:

a. El número de elementos está completamente delimitado

b. El número de elementos está escondido

c. El número de elementos es mayor que 1.000

d. Ninguna de las anteriores

e. El número de elementos no está limitado

2. Para determinar que una buena muestra seleccionada del sector industrial es la correcta esta debería ser:

a. Pequeña

b. Grande

c. Mediana

c. Representativa de la población objetivo

d. e. Incluir sólo establecimientos pequeños

3. Se selecciona una muestra es aleatoria cuando:

a. En forma repetitiva

b. Por conveniencia

c. Al azar

d. De manera que todas tengan la misma posibilidad

e. A través de un censo

4. Una medida aplicada a la característica de la unidad de una población se llama:

a. Parámetro

b. Estimador

c. Variable

d. Dato

e. Población

5. Un ejemplo de característica cualitativa puede ser:

a. Gasto mensual en alimentación

b. Pulsaciones por minuto

c. Salarios

d. Ocupación

e. Todas las anteriores

Ejercicios Propuestos

1. Con los siguientes datos identificar y clasificar según el tipo de dato al cual pertenezca, ya sea cuantitativos (numéricos) y cualitativos (categóricos), o discretos o continuos:

    a. Peso de kilogramos de 19 agricultores.
    b. Promedio de notas en los seguimientos trimestrales de una planta.
    c. Escolaridad de 189 empleados
    d. Numero de acero utilizado en estructuras metálicas.
    e. Trayecto recorrido en kilómetros por un automóvil durante 1 año
    f. Dinero ahorrado por 5 personas por un periodo de tiempo.
    g. Estatura de un grupo de personas.
    h. Promedio de dureza en un material

2. Explique y de ejemplo qué entiende usted por:

a. Marca de clase.
b. Variable continúa.
c. Intervalo de clase.
d. Histograma de frecuencias.
e. Polígono de frecuencias.
f. Amplitud del intervalo.

3. Responder verdadero o falso a las siguientes afirmaciones, en caso de que sea falso, cambiar la palabra **subrayada** por otra, en la cual el enunciado tenga validez.

a. En un grupo de valores se le puede calcular **más de un** promedio.
b. En una serie de datos, cuando el valor es **par**, la mediana es igual al valor central.
c. La **moda** divide a los datos de una distribución en dos mitades: la mitad mayores y la mitad menores que su propio valor.
d. La **media** es el valor de la variable que más aparece en los datos de una distribución.

4. Responder verdadero o falso a las siguientes afirmaciones**,** justificar la respuesta.

a. Si la desviación estándar de un conjunto de datos es 0, entonces, los datos son iguales.
c. Existen datos con desviación estándar negativa.
d. En una distribución simétrica, la media, la mediana y la moda son iguales.
e. La desviación estándar esta dada por las mismas unidades que la media.

f. Toda información numérica proporciona datos cuantitativos.

g. Toda información no numérica ofrece datos cuantitativos.

h. Cuando todos los datos son categóricos, la moda es la única medida de tendencia central que se puede utilizar.

i. Si el primer cuartil en el primer examen de estadística fue de 3,0, entonces, este valor indica que el 25% de los estudiantes ganaron el examen.

6. Los siguientes datos representan los ahorros mensuales en miles de pesos, de 25 trabajadores en una fábrica de guante.

21,48	21,15	25,12	23,47	27,81	19,8	36,05	28,5	26,66
20,35	30,22	25,49	20,8	23,83	25,35	23,48	25,81	21,07
26,83	30,96	33,38	20,77	19,98	35,87	22,02		

a. ¿Qué porcentaje del grupo pagó menos de 21.000 pesos?

b. ¿Qué porcentaje pagó más de 22.000 pesos pero menos de 27.000 pesos?

7. En una empresa de hilos recubiertos se desea saber cuál es el grado de escolaridad de las personas que trabajan en la misma, sabiendo que la empresa cuenta con una población de 800 personas, se extrajo una muestra de 80 personas de diferentes áreas de la compañía.

8. La escolaridad de la muestra tomada es la siguiente:

Primaria: 22
Secundaria: 30
Técnicos: 15
Tecnólogos: 8
Universitarios: 5

a) Hallar cual es la frecuencia absoluta
b) Hallar cual es la frecuencia acumulada
c) Hallar cual es la frecuencia relativa
d) Crear un diagrama de frecuencias e interpretar

9. En una población de estudiantes universitarios de ingeniería, se seleccionan al azar una muestra de 15 alumnos por el total de estudiantes existentes (534 estudiantes), y se quieren determinar sus promedios por el numero de materia que posen (5 materias)

5	4	1	4	5	1	4	3	3	2
4	1	4	5	5	4	4	3	2	4
3	4	3	3	4	5	4	1	2	4
2	3	2	4	5	1	4	4	2	4
2	1	4	3	5	3	2	5	5	3
3	4	1	5	1	5	4	4	4	2
2	5	4	4	3	4	1	4	4	5
2	5	4	2	4					

En la universidad se poseen unos criterios de evaluación

a. Crear un histograma de frecuencias en el que se identifiquen cada una de las clases.
b. Hallar cual es la frecuencia acumulada.
c. Hallar cual es la frecuencia relativa

10. La siguiente información muestra el pero en kilogramos de 50 piezas tomadas de una máquina de hacer poleas y estos son

83	51	66	61	82	65	54	56	92	60
65	87	68	64	51	70	75	66	74	68
44	55	78	69	98	67	82	77	79	62
38	88	76	99	84	47	60	42	66	74
91	71	83	80	68	65	51	56	73	55

Use datos para construir las distribuciones de frecuencia relativa con 7 y 13 intervalos iguales. Se debe de tener más de 50% del peso mayor al 50Kg, de lo contrario las piezas se tomaran como no conformes.

- **a.** ¿Cumplen las presas con lo pedido?
- **b.** ¿La distribución de frecuencias relativas de 13 intervalos ayuda a responder el inciso a) mejor que la distribución de 7 intervalos?
- **c.** Se desea saber que porcentaje de los pesos se encuentran entre 45 y 50 Kg de edad. ¿A partir de cuál distribución de frecuencias relativas, de 7 o de 13 intervalos, puede estimar la mejor respuesta?
- **d.** Calcule la media, la mediana y la moda para 7 y 13 intervalos y determinar el porcentaje de error con los datos hallados en el STAGRAPHICS.
- **e.** Calcule la varianza, desviación estándar, coeficiente de variación para 7 y 13 intervalos y determine el porcentaje de error con los datos hallados.
- **f.** Para los 7 y 13 intervalos hacer los diagramas de frecuencia, la ojiva y la grafica de torta para intervalos.

10. En una clínica se han llevado 50 muestras de sangre para una investigación, en las muestras observadas se mira que uno de los componentes de la sangre se tiene en las siguientes cantidades:

25.3	20.7	22.5	21.2	23.8	23.3	20.9	22.9	23.5	19.5
20.8	22.8	21.9	22.0	20.7	20.9	25.0	22.2	22.8	20.1
23.7	20.3	23.6	19.0	25.1	25.0	19.5	24.1	24.2	21.8
19.7	24.2	23.8	20.7	23.8	24.3	21.1	20.9	21.6	22.7
21.3	21.5	23.1	19.9	24.2	24.1	19.8	23.9	22.8	23.9

a. Ordenar los datos en un arreglo ascendente.

b. Construya una distribución de frecuencias absolutas y una distribución de frecuencias acumuladas "menor que" a partir de los datos. Utilice intervalos de 0,8 minutos.

c. Construya un polígono de frecuencia con base en los datos.

d. A parir de los datos, construya una Ojiva "menor que".

e. Tomando en cuenta su Ojiva, estime qué porcentaje de sustancia existe en la sangre que sea inferior a 24 ml.

11. Los trabajadores de una empresa solicitan en una convención colectiva que cada salario semanal de sus afiliados sea aumentado según la ecuación:

$$y_i = 2{,}27x_i + 67.750$$

La empresa tiene 9.560 trabajadores y antes de solicitar el reajuste salarial devengaban un promedio de $496.000 mensual.

a. ¿Cuál será el nuevo promedio de ingreso trimestral de los trabajadores si la empresa acepta?

b. El coeficiente de variación antes de solicitar el reajuste salarial es del 25%. ¿Cuál será el nuevo coeficiente de variación?

12. En un grupo de 120 personas, se sabe que 70 de ellas tienen un salario medio diario de $17.000 los restantes de $36.000.

a. Se pide calcular la ganancia media de 120 personas.

b. Si se sabe que el coeficiente de variación es del 35%, ¿Cuál es la varianza de los salarios diarios de las 120 personas?

13. En una empresa de plásticos, los jornales semanales tienen una media de $246.000. Como una solución al conflicto laboral surgido se proponen dos soluciones al conflicto:

- **a.** un aumento del 3,5% en el salario semanal
- **b.** un aumento del 6%, más una bonificación semanal de $10.000 por cada obrero
- **c.** ¿cuál de las dos alternativas mejora la situación del obrero?

14. En una fábrica de hilos recubiertos se proyecta lanzar al mercado un nuevo tipo de hilo que cumple la función del chaleco antibalas. Se realiza un test de aceptación en una muestra de 40 políticos utilizando una escala de 10 puntos. Los valores obtenidos fueron los siguientes:

2 • 6 • 8 • 7 • 4 • 5 • 10 • 6 • 6 • 7 • 6 • 7 • 3 • 8 • 7 • 6 • 8 • 6 • 5 • 4 • 7 • 8 • 5 • 7 • 6 • 7 • 2 • 7 • 2 • 7 • 4 • 6 • 7 • 1 • 8 • 7 • 9 • 1 • 6 • 9

Hacer una tabla de frecuencia en donde se encuentre:

- **a.** Marca de clase
- **b.** Intervalos
- **c.** Frecuencia (Relativa, Absoluta, Acumulada)
- **d.** Rango

ESTUDIO DE CASOS

CASO 1

El observatorio del Caribe colombiano, la Universidad del Atlántico y la Cámara de Comercio de Cartagena, con el apoyo del Grupo Bancolombia, presentan un nuevo número de la serie de estudios sobre la competitividad de Cartagena, en este caso analiza la situación competitiva de 22 departamentos de Colombia y la ciudad de Bogotá D.C., centrando específicamente el análisis en los departamentos de Atlántico y Bolívar.

La siguiente tabla muestra las fortalezas y debilidades en competitividad. Se entienden como fortalezas los indicadores en los cuales las ciudades se hayan ubicado en los 10 primeros lugares, y debilidades aquellos en los que ocuparon las 10 últimas posiciones.

Tabla 2.20. FORTALEZAS Y DEBILIDADES EN COMPETITIVIDAD, 2008

DEPARTAMENTO	FORTALEZAS	DEBILIDADES	BALANCE
Bogotá	45	6	39
Santander	44	6	38
Atlántico	38	4	34
Antioquia	38	7	31
Valle del Cauca	39	9	30
Caldas	30	12	18
Cundinamarca	31	14	17
Risaralda	30	15	15
Bolívar	23	18	5
Boyacá	22	20	2
Norte de Santander	20	21	-1
Meta	22	26	-4
Quindío	25	19	-6
Cesar	18	30	-12
Tolima	10	22	-12
Cauca	15	30	-15

Huila	8	25	-17
Nariño	13	14	-21
Sucre	12	37	-25
Córdoba	9	38	-29
La Guajira	10	39	-29
Magdalena	6	35	-29
Caquetá	2	43	-41

Fuente: Estudios sobre la competitividad de Cartagena, 2008

De esta forma, se resaltan las ventajas y desventajas de los departamentos, con el fin de hacer un balance que nos muestre la situación competitiva, conociendo de manera precisa aquellos indicadores pasivos y activos que influyen en la competitividad de los departamentos.

a) Realice un análisis estadístico de los datos, donde muestre el grafico adecuado para la representación de los datos, explicando el porqué de la escogencia de este grafico y cuál era la falencia de los otros gráficos.

b) El DANE quiere saber si los datos obtenidos presentan algún tipo de tendencia o se ajustan algún tipo de distribución.

CASO 2.

MOVILIDAD EN TRANSPORTE PÚBLICO COLECTIVO EN CARTAGENA[6]

El proyecto Cartagena Cómo Vamos (CCV) llevó a cabo un monitoreo mediante trabajo de campo, de los tiempos y velocidades promedio de desplazamiento del transporte público en la ciudad de Cartagena.

La herramienta aplicada fue diseñada en septiembre de 2005 para CCV por un equipo de ingenieros civiles y de transporte, economistas y estadísticos de las Universidades de los Andes, Nacional, Cartagena y Tecnológica de Bolívar, y desde entonces es aplicada semestralmente con el apoyo de profesores y estudiantes de la Universidad Tecnológica de Bolívar, y en esta oportunidad con el apoyo de la Facultad de Ingeniería Civil de la Universidad de Cartagena.

Hay que tener en cuenta que en el trabajo de campo se miden los tiempos y velocidades de desplazamiento en las cinco rutas de transporte público colectivo que movilizan el 37% de los pasajeros según un estudio de la Universidad Nacional de 2002: bus de Ternera-Centro-Laguito (TCL), bus de Socorro-Bosque-Manga (SBM), buseta de Bosque (B), buseta de Socorro-Jardines (SJ), y buseta de Ternera-Avenida (TA). Estas rutas cubren la Avenida Pedro de Heredia y la Avenida del Bosque que son las dos arterias viales de la ciudad. Las muestras son tomadas semestralmente en ambos sentidos de las rutas (del centro a la periferia y de la periferia al centro) durante la mañana, mediodía y tarde y en semanas sin festivos. Debido a la ejecución de las obras de TRANSCARIBE, algunas partes del recorrido de las rutas evaluadas han sufrido variaciones. En la prueba efectuada en el primer semestre del año 2006 se registraron cambios con respecto a la ruta original, y en el segundo semestre de 2008 se volvieron a registrar otros. Estos

[6] Tomado y adaptado de Cartagena como Vamos.

cambios incidieron en la distancia de algunos recorridos. En el primer semestre de 2009 no hubo cambios.

En la última medición realizada entre mayo y junio de 2009, no se registraron cambios frente a los resultados de los cuatro años anteriores
Los siguientes datos son los relacionados con el tiempo promedio de desplazamiento entre las diferentes horas picos y las horas no pico:

Tabla 2.21. Tiempo promedio de desplazamiento

Tiempo promedio de desplazamiento	Cualquier Franja Horaria (min)
primer periodo de 2006	53
Segundo periodo de 2006	64
primer periodo de 2007	65
Segundo periodo de 2007	62
primer periodo de 2008	61
Segundo periodo de 2008	62
primer periodo de 2009	58

Fuente: Proyecto Cartagena Como Vamos

Tabla 2.22. Velocidad promedio de desplazamiento

Velocidad promedio de desplazamiento	Cualquier Franja Horaria (Km/hr)
Segundo periodo de 2005	15.4
primer periodo de 2006	15.6
Segundo periodo de 2006	15.2
primer periodo de 2007	15.7
Segundo periodo de 2007	15.2
primer periodo de 2008	15.6
Segundo periodo de 2008	16
primer periodo de 2009	16.6

Fuente: Proyecto Cartagena Como Vamos

Para el año 2009, se obtuvieron estos tiempos y velocidades relacionados con la movilidad en cada una de las rutas:

Tabla 2.23. Tiempo Promedio en horas pico

Tiempo	Mañana	Medio día	Tarde
Global	58,2	63,1	64,0
TCL	66,1	71,1	69,7
SBM	64,4	68,6	74,1
B	56,4	59,2	64,4
SJ	52,2	59,0	53,7
TA	52,0	57,9	58,3

Fuente: Proyecto Cartagena Como Vamos

Tabla 2.24. Velocidad Promedio en horas pico

Velocidad	Mañana	Medio Día	Tarde
Global	15,7	15,6	16,1
TCL	15,2	14,7	15,7
SBM	18,9	18,9	17,9
B	16,8	18,1	17,4
SJ	13,3	12,9	15,2
TA	14,5	13,5	14,1

Fuente: Proyecto Cartagena Como Vamos

TCL = Ternera-Centro-Laguito

SMB = Socorro-Manga-Bosque

B = Bosque

SJ = Socorro-Jardines

TA = Ternera-Avenida

Tabla 2.25. Tiempo Promedio en horas no pico

Tiempo	Mañana	Medio día	Tarde
Global	62	64	66
TCL	71	69	73
SBM	66	68	75
B	62	64	67
SJ	55	57	59
TA	54	62	57

Fuente: Proyecto Cartagena Como Vamos

Tabla 2.26. Velocidad Promedio en horas no pico

Velocidad	Mañana	Medio día	Tarde
Global	15	15	15
TCL	14	16	15
SBM	19	18	18
B	14	14	14
SJ	14	13	14
TA	13	13	14

Fuente: Proyecto Cartagena Como Vamos

Tabla 2.27. Tiempo Promedio de acuerdo a franja horaria

Ruta Pedro Heredia		Ruta Bosque	
Franja Horaria	Tiempo	**Franja Horaria**	Tiempo
Global	62	**Global**	67
Pico	63	**Pico**	68
No Pico	60	**No Pico**	64

Fuente: Proyecto Cartagena Como Vamos

Tabla 2.28. Velocidad Promedio de acuerdo a franja horaria

Ruta Pedro Heredia		Ruta Bosque	
Franja Horaria	Velocidad	**Franja Horaria**	Velocidad
Global	14	**Global**	16
Pico	14	**Pico**	16
No Pico	15	**No Pico**	17

Fuente: Proyecto Cartagena Como Vamos

1. Realice un paper donde se muestre el análisis y gráficos estadísticos de las variables: tiempo y velocidad. Además determine si existe alguna relación de dependencia entre éstas variables, el tipo de relación y las consecuencias de esto.

2. Con base en el análisis estadístico, determinar cuál es el mayor y menor tiempo y velocidad que puede existir.

3. Teniendo en cuenta las variables Tiempo y Velocidad. Además si usted puede escoger el horario en el cual desplazarse, ¿cuál de las rutas escogería en el horario de la mañana, medio día, y tarde?

4. De acuerdo a la información suministrada, ¿Cuál ruta (Pedro de Heredia o Bosque) es más conveniente tomar en horario pico, teniendo en cuenta el tiempo y la velocidad?.
5. Suponiendo que para el 2012 la velocidad promedio del transporte público colectivo será de 28 km/hr. Determinar cuál es la ganancia en tiempo del implementar un sistema masivo de transporte transcaribe. ¿Cuál es la ganancia en minutos de que se ahorraría en implementar este sistema?

CAPITULO 3

TEORIA DE LA PROBABILIDAD

INTRODUCCION

El azar está presente en la vida cotidiana en muchos contextos en los que aparecen nociones de incertidumbre, riesgo y probabilidad, en los que la persona debe tomar decisiones que le pueden afectar, emitir juicios sobre relación entre sucesos o efectuar inferencias y predicciones (Gigerenzer, 2002).

La teoría de la probabilidad proporciona métodos para cuantificar las oportunidades, o, probabilidades asociadas con varios resultados. En la actualidad, la teoría de la probabilidad constituye el fundamento de las aplicaciones estadísticas en diferentes áreas como la investigación científica, economía, psicología, sociología, ingeniería y otras como base para la toma de decisiones.

La probabilidad inicialmente surgió como un concepto matemático, pero actualmente recibe múltiples significados que se deben principalmente al tipo de problemas que se resuelven, la forma de asignar probabilidades e incluso sus propiedades.

El desarrollo de la teoría de la probabilidad ha aumentado el alcance de las aplicaciones de la estadística. Muchos conjuntos de datos se pueden aproximar, con gran exactitud, utilizando determinadas distribuciones probabilísticas; los resultados de éstas se pueden utilizar para analizar datos estadísticos. La probabilidad es útil para comprobar la fiabilidad de las inferencias estadísticas y para predecir el tipo y la cantidad de datos necesarios en un determinado estudio estadístico.

En este capítulo se presenta el bosquejo histórico de la teoría de la probabilidad, los conceptos fundamentales de la teoría de probabilidad tales como experimentos aleatorios, eventos y espacio muestral. Además las técnicas de conteo, probabilidades condicionales e independientes.

COMPETENCIAS:

- Conjeturo acerca del resultado de un experimento aleatorio usando proporcionalidad y nociones básicas de probabilidad.

- Resolucion y formulacion de problemas a partir de un conjunto de datos presentados en tablas, diagramas de barras, diagramas circulares.

- Prediccion y justificacion de razonamientos y conclusiones usando información estadística.

Comparacion de resultados de experimentos aleatorios con los resultados previstos por un modelo matemático probabilístico.

Utilizacion comprensivamente de algunas medidas de centralización, localización, dispersión y correlación (percentiles, cuartiles, centralidad, distancia, rango, varianza, covarianza y normalidad).

3. LA TEORÍA DE LA PROBABILIDAD

3.1. HISTORIA

La probabilidad ha existido en formas sencillas, desde el inicio de las civilizaciones en Grecia, Roma, e India explicaban el azar mediante la voluntad divina, utilizaban la configuración resultante de tirar cuatro dados para predecir el futuro y revelar la voluntad favorable o desfavorable de los dioses.

En el renacimiento, surge un nuevo enfoque global de considerar al mundo en el cual el abandono progresivo de las explicaciones teológicas conduce a una reconsideración de los experimentos aleatorios. Los matemáticos italianos del siglo XVI, comienzan a interpretar los resultados de experimentos aleatorios simples. Cardano, establece la equiprobabilidad de aparición de las caras de un dado a largo plazo. A finales del siglo XVI, existía un intuitivo pero preciso análisis empírico de los resultados aleatorios.

Durante los siglos XVI y XVII se desarrolla el análisis matemático de los juegos de azar y algunos autores consideran como origen del cálculo de probabilidades la resolución del problema de los puntos en la correspondencia entre Pascal y Fermat en 1654.

Pero es tan solo a comienzos del siglo XVIII cuando la probabilidad se consolida como disciplina. Posteriormente, el cálculo de probabilidades se extiende a problemas físicos debido a los problemas de astronomía y física que surgen ligados a la contrastación empírica de la teoría de Newton. Estas investigaciones van a ser de importancia fundamental en el desarrollo de la Estadística.

La teoría de la probabilidad fue aplicada en gran medida en los juegos de azar, por medio de la teoría de juegos, que tuvo sus inicios con una publicación de Von Neumann titulada "The Game Theory and Economic Behaviour", la teoría de juegos muestra una solución para los juegos de suma cero y con fundamentos para el análisis de juego entre más de dos jugadores, después aportó Nash soluciones para una gama amplia de juegos.

Ante los problemas socioeconómicos y el surgimiento de la industria de los seguros en el siglo XIX, se inició la enseñanza de la probabilidad como un instrumento que les permitiría entender los fenómenos sociales y conocer con exactitud el riesgo de perder a través del cálculo de las pólizas de seguros.

D. Bernoulli fue pionero en la aplicación del cálculo infinitesimal al cálculo de probabilidades, proporciona la primera solución aplicando la teoría de errores, al estimar una cantidad desconocida a partir de un conjunto de mediciones de su valor que por el error experimental presentan variabilidad.

Es importante destacar autores como Abraham de Moivre, el reverendo Thomas Bayes y Joseph LaGrange quienes inventaron fórmulas y técnicas de probabilidad.

El impulso fundamental de la teoría de las probabilidades proviene de la obra de Pierre Simon, Marqués de Laplace, quien indujo la primera definición explícita de probabilidad y desarrolló la ley normal como modelo para describir la variabilidad de los errores de medida; también formuló y estimó el primer modelo explicativo estadístico. Por su parte, Gauss hizo su aportación en la estimación de modelos estadísticos.

Bravais, realizo su aporte al considerar la relación entre errores de medida dependientes entre sí; Benjamín Pierce propone el primer criterio para rechazar observaciones heterogéneas con el resto y S. Newcomb, el más famoso astrónomo americano del siglo XIX, introduce los primeros métodos de estimación cuando hay errores fuertes en algunos datos (Estimación Robusta).

A finales del siglo XIX y en todo el siglo XX, se dio lugar a la creación de diferentes escuelas y tendencias dedicadas al estudio de la teoría de la probabilidad, las más importantes son:

- La escuela rusa que dominó todas las áreas relativas al cálculo de probabilidades y la estadística, la escuela fue dirigida por Andrei N.Kolmogorov (1903-1987) y Khinchine. Entre sus predecesores se encontraron Chebyshev, Markov y Liapunov entre otros.
- La escuela estadounidense: Los principales exponentes de esta escuela fueron Feller y Doob, aunque el iniciador de este movimiento fue Nortber Wiener (1894-1964) quien desarrolló una medida de las probabilidades para conjuntos de trayectorias que no son diferenciables en ningún punto, asociando una probabilidad a cada conjunto de trayectorias. William Feller se hizo conocido por sus estudios relacionados con el Teorema Central del Límite, además de la demostración de la condición de Lindeberg.
- La escuela francesa: Entre sus exponentes se encuentran Meyer, Nevev y Fortret de París, entre sus aportes más importantes se encuentra la generalización del concepto de diferenciación utilizando la teoría de distribuciones.

3.2. DEFINICIÓN DE LA PROBABILIDAD

El término probabilidad se refiere al estudio de la aleatoriedad y la incertidumbre, en cualquier situación donde se produzca alguno de varios resultados posibles. La probabilidad mide la frecuencia con la que ocurre un resultado en un experimento bajo condiciones suficientemente estables. Además se puede considerar que la probabilidad es la percepción o grado de creencia que las personas tienen de la ocurrencia de un suceso.

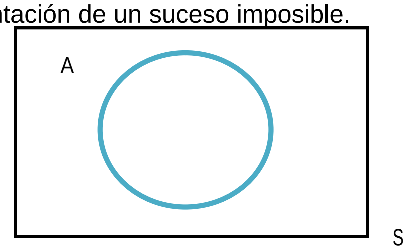
$$= \frac{\bullet\bullet\ \bullet\ \bullet\bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet}{\bullet\bullet\ \bullet\ \bullet\bullet\bullet\ \bullet\bullet}$$

La probabilidad se puede definir como la frecuencia relativa de aciertos o casos favorables. Es decir, la ocurrencia de un evento determinado en un proceso aleatorio en el cual se ha repetido un gran número de pruebas o experimentos. La frecuencia relativa es el número de casos favorables dividido entre el número de pruebas efectuadas

Todas las probabilidades deben satisfacer los siguientes axiomas:[7]
Estos axiomas se pueden interpretar como una regla para el cálculo de las probabilidades:

- La probabilidad cuando se presenta en un suceso imposible es 0: P(∅)=0

Figura 3.1. Representación de un suceso imposible.

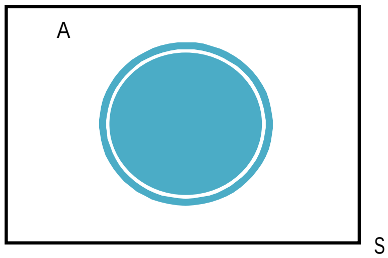


Fuente: Los Autores

[7] *Tomado y Adaptado de Curso De Especialista En Gestión y Control De La Calidad Departamento De Estadística E Investigación Operativa*, *Universidad Politécnica de Valencia, Autores Andrés Carrión García Mª Teresa Carot Sánchez*

- La probabilidad cuando se presenta en un suceso seguro es 1.

Figura 3.2. Representación de un suceso seguro.

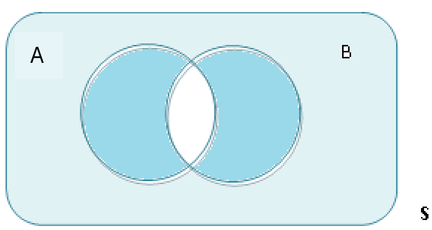


Fuente: Los Autores

- Una probabilidad nunca puede ser menor que 0 ni mayor que 1: : La probabilidad comprende números entre y . Desde otro punto de vista se puede también decir la probabilidad como la *proporción* de muestra en los que se verifica una acción.

- La probabilidad de la unión de dos sucesos es:

Figura 3.3. Representación de la unión de sucesos.



Fuente: Los Autores

Y en general podemos decir que:

$$P\left(\bigcup_i A_I\right)=\sum_i P(A_i)+\sum_i\sum_j P(A_i\cap A_j)+\cdots+(-1)^{n-1}P\left(\bigcap_i A_i\right) \qquad \text{i=1,2,...,n}$$

Las probabilidades se pueden estimar mediante:

- La frecuencia relativa de eventos pasados: Cuando se realiza un experimento controlado o un muestreo en una población finita, las probabilidades se deduce de la experiencia obtenida de la observación prolongada.

Ejemplo 3.1:

En un estudio sociológico se determino que la mayoría de los casos de violencia intrafamiliar se da porque el padre puede poseer problemas con el alcohol, es desempleado o sufrió de maltrato en su infancia, se desea calcular la probabilidad de que los encuestados sean personas que no hayan sufrido algunos de los tres problemas anteriores y tengan problemas de violencia intrafamiliar:

Para determinar esto se aplicó una encuesta a 30 hombres y los resultados obtenidos fueron:

Tabla 3.1. Casos de violencia intrafamiliar

Encuestados	Frecuencia
Alcohólicos, Desempleados, Maltrato de en su infancia	18
Violencia Intrafamiliar sin ningún suceso anterior	12
Sumatoria	30

Fuente: Los Autores

Siendo (•) = Personas Alcohólicas, desempleados, y que sufrieron algún maltrato en su infancia

Siendo (1 • •) = Personas que sufren de violencia intrafamiliar sin ningún suceso anterior (Alcoholismo, desempleo, maltrato en la infancia)

$$\bullet(\bullet) = \frac{\bullet\ \bullet\ \bullet\ \bullet\bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet \qquad \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet\bullet\bullet\ \bullet}{\bullet\text{ú}\ \bullet\ \bullet\bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet\ \bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\ \bullet\bullet\bullet} = \frac{\bullet(\bullet)}{\bullet(\bullet)} = \frac{18}{30} = 0.6$$

Por lo tanto, la probabilidad de que los encuestados sean personas que no hayan sufrido algunos de problemas de Alcoholismo, desempleo y maltrato en la infancia y tengan problemas de violencia intrafamiliar es $\bullet(1\ \bullet\ \bullet) = 0.4$

- Distribuciones teóricas: Las probabilidades pueden determinarse sin recurrir a las frecuencias relativas. Para esto se utilizan distribuciones de probabilidad como la binomial, en la cual la validez de dichas distribuciones teóricas depende del grado que la hipótesis representen la realidad.

- Apreciación subjetiva: Muchas veces la toma de decisiones dependen del juicio o criterio de una persona con experiencia, el cual se basa en esta experiencia para estimar las probabilidades. Una probabilidad subjetiva es una evaluación que una persona que toma decisiones hace acerca de la ocurrencia de un evento incierto. La probabilidad estimada es de carácter personal, por lo tanto se puede encontrar que un mismo evento tenga diferentes probabilidades de ocurrencia.

Clases de probabilidad entre las cuales tenemos:

Figura 3.4. Tipos de probabilidades.

Probabilidad a Priori

- Es aquella que se puede definir de manera inmediata, sin necesidad de realizar pruebas o experimentos.
- Ejemplo: la probabilidad de que salga un 6 en el lanzamiento de un dado, se sabe que es igual a 0,16666

Probabilidad Empirica

- Este es determinada por pruebas y experimentos
- Ejemplo: la probabilidad de ganar un partido de futbol: será: número de partidos ganados, entre el total de partidos
- P=# de casos favorables/# de casos posibles

Fuente: Los Autores

Ejemplo 3.2.

En un estudio psicológico la relación entre las condiciones psiquiátricas y el comportamiento de una persona apostadora, ya sea joven o adultos. El estudio fue realizado a 16 jóvenes y 25 adultos. Se selecciona al azar una persona de esa muestra. Se desea saber cuál es la probabilidad de que un joven se comporte como un apostador.

$$P(A) = \frac{16}{41} = 0{,}39$$

3.3 EXPERIMENTOS ALEATORIOS

Un experimento aleatorio es un fenómeno que puede arrojar diferentes resultados, no siendo previsible los mismos, debido a la incertidumbre con la cual se producen dichos resultados ya sea a partir de la observación o experimentación.

Características de un experimento aleatorio:

- Un experimento se puede repetir de manera indefinida en iguales condiciones
- Cuando se modifican las condiciones iniciales y se repetirá la experimentación, los resultados pueden variar
- Se pueden establecer un conjunto de posibles resultados pero no pronosticar un resultado particular
- Se pueden obtener un modelo estadístico de los datos siempre y cuando el experimento se repita un número considerable de veces.

3.4. ESPACIO MUESTRAL:

Es el conjunto de los resultados posibles, los cuales se obtienen de realizar un experimento aleatorio o un muestreo de tipo aleatorio, se representa por •, existen diferentes tipos de espacios muéstrales entre los cuales está:

- Espacio muestral discreto: Cuando el conjunto está formado por elementos finitos o infinitos de números enteros.

  Ejemplo 3.3:

  Juan se dirige en su auto por la avenida Pedro de Heredia, él está pensando en el semáforo que se encuentra en el sector cuatro vientos, el desea encontrar la luz en verde, pero esto no se sabe con certeza porque podría encontrar el semáforo con la luz en rojo ó amarillo.

  Por lo cual, el espacio muestral (•) asociado al experimento aleatorio es:

  • = {•••••, ••••, • • ••••••}

Ejemplo 3.4:

En una empresa que fabrican refrigeradores, realizan un estricto control de calidad de sus productos, para esto seleccionan una muestra de cada lote de fabricación, la muestra es igual a 3 refrigeradores. Cuando se realiza la prueba de calidad, el espacio muestral que contiene los posibles resultados de encontrar productos defectuosos en un lote de fabricación es:

$$\bullet = \{0, 1, 2, 3\}$$

- Espacio muestral continuo: Es el conjunto de resultados, los cuales son intervalos de tipo continuo, es decir que entre dos resultados existen infinitos de resultados posibles, por ejemplo:

  La selección de un punto cualquiera al azar de un punto del segmento [0,1],que está formado por el par (E, U), donde tenemos que:

$$\bullet = \{\bullet \ \bullet \ \bullet | 0 ^{\bullet} \bullet \ \bullet \ \bullet \ 1\}$$

3.5. SUCESO O EVENTO ALEATORIO

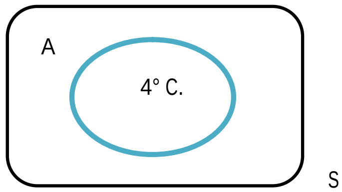
Un suceso es un subconjunto de resultados elementales de un experimento aleatorio y que pertenecen al espacio muestral, existen diferentes tipos de sucesos o eventos, los cuales son:

- ***Suceso elemental:*** También conocido como suceso simple o punto muestral, se da en el momento que el suceso se encuentra formado por un único elemento del espacio muestral

  Ejemplo 3.5:

  La medición de la temperatura de un cuarto frío en horas específicas del día, cada una de las temperaturas que se obtienen representan un elemento.

Figura 3.5. Diagrama de Venn de un suceso elemental.



Fuente: Los Autores

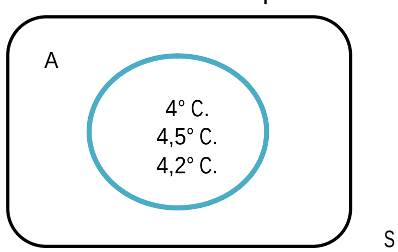
Si el espacio muestral es la temperatura que se toman, la temperatura a la cual se encuentra el cuarto frio a la 1:00 p.m. corresponde al suceso elemental

- ***Suceso compuesto:*** Es cuando el suceso se encuentra formado por más de un elemento del espacio muestral.

Ejemplo 3.6:

Si la medición de la temperatura de un cuarto frio se realiza durante una hora, las temperaturas que se obtienen son los sucesos elementales.

Figura 3.6. Diagrama de Venn de un suceso compuesto.



Fuente: Los Autores

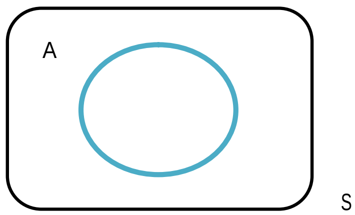
Si el espacio muestral es la temperatura que se toman, la temperatura en la que se encuentra el cuarto frio entre la 1:00 p.m. y 2:00 p.m. corresponde al conjunto de sucesos.

- ***Suceso imposible:*** Es el suceso que nunca ocurre. Esto es, • • •

Ejemplo 3.7:

Al lanzar un dado se obtenga 7 como resultado.

Figura 3.7. Diagrama de Venn de un suceso imposible.



Fuente: Los Autores

- ***Suceso seguro:*** Es el suceso que siempre ocurre, por lo que existe un espacio muestral constituido por todos los sucesos elementales.

Ejemplo 3.8:

María tiene una bolsa que contiene canicas de color azul y Luis extrae una canica azul. Cualquier canica que Luis desee extraer será de color azul.

Figura 3.8. Diagrama de Venn de un suceso seguro.

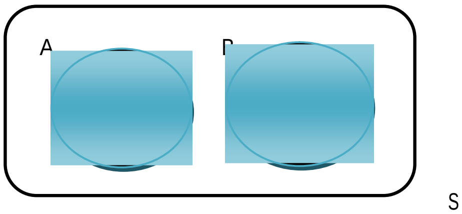
Fuente: Los Autores

- ***Sucesos mutuamente excluyentes:*** Son aquellos que no se pueden presentar de manera simultánea. Es decir, • • • = •

Ejemplo 3.9:

En McDonald hay dos colas, la primera cola es para realizar el pedido de comidas rápidas y la segunda cola es para realizar el pedido de productos de heladería. Cuando un cliente llega a McDonald tiene la opción de unirse a la cola para realizar el pedido de comidas rápidas o unirse a la cola para realizar el pedido de productos de heladería, pero no las dos al mismo tiempo.

Figura 3.9. Diagrama de Venn de sucesos mutuamente excluyentes



Fuente: Los Autores

- ***Suceso dependiente:*** Este aquel donde el suceso que se presente depende de un suceso anterior o de otro suceso.

Ejemplo 3.10:

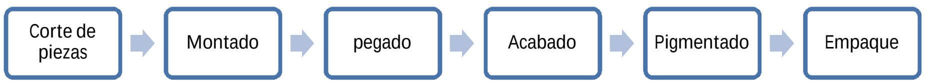
En el proceso de fabricación de un calzado se realizan las siguientes actividades:

1. Corte de piezas: Se realiza el corte de la moldura de acuerdo a las medidas y el modelo diseñado.
2. Montado: Se selecciona la horma de acuerdo a la numeración para conformar y fijar la planta a base de clavos.
3. Pegado: Se hace una hendidura para que el pegamento se impregne mejor y posteriormente se realiza el pegado de la suela
4. Acabado: Se pega la plantilla, se pintan los bordes de la suela y forros, se desmancha el zapato de residuos del proceso productivo.
5. Pigmentado: Se realiza con el fin de uniformizar el color.

6. Empaque: Se guarda el producto en caja de cartón

Para realizar la segunda actividad es necesario realizar la primera, gráficamente:

Figura 3.10. Proceso de fabricación de calzado



Fuente: Los Autores

- ***Suceso compatible:*** Son aquellos sucesos que se pueden presentar de manera simultánea.

Ejemplo 3.11:

Para fabricar masilla, el operario debe vaciar en la mezcladora las materias primas para iniciar el proceso de mezclado, en este tiempo un segundo operario realiza el alistamiento de los envases donde será vaciada la masilla. Sea:

•: vaciar las materias primas en la mezcladora

•: Alistar los envases

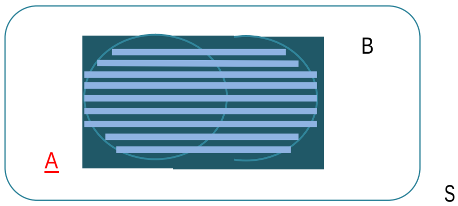
Estas dos actividades se realizan en el mismo tiempo, por lo cual son compatibles.

3.5.1. OPERACIONES CON EVENTOS:

Teniendo en cuenta la teoría de conjuntos tratada en el primer capítulo, se pueden explicar las operaciones con sucesos.

- **Unión:** (• • •) Subconjunto de elementos de • que están incluidos, al menos en uno de esos eventos (A o B)

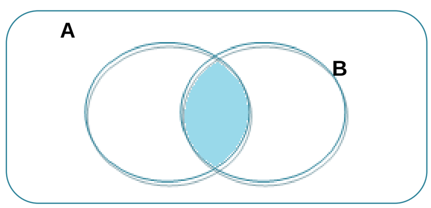
Figura 3.11. Diagrama de Venn de unión de eventos.



Fuente: Los Autores

- **Intersección:** Es el subconjunto que se encuentran incluidos al mismo tiempo en los subconjuntos de ambos eventos.

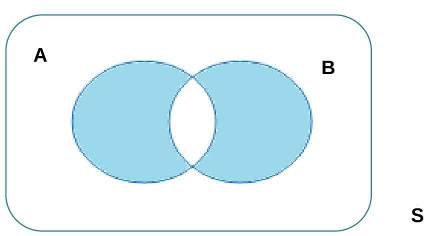
Figura 3.12. Diagrama de Venn de intersección de eventos.



Fuente: Los Autores

Diferencia de eventos: es el subconjunto de sucesos formados por todos los elementos que son parte de pero no de . esto es:

Figura 3.13. Diagrama de Venn de diferencia de sucesos.



Fuente: Los Autores

Ejemplo 3.12:

“Se lanza un dado y se observa el número de puntos que aparecen en sus lados”, este puede presentar seis posibles resultados; el espacio muestral de este es

• = {1, 2, 3, 4, 5, 6} y • = 6

Un resultado posible de un lanzamiento es que salga un número impar. Este tipo de resultado podemos llamarlo como evento y en este caso seria *A= {1, 3, 5}* y seria e=3

3.6. TECNICAS DE CONTEO

Las técnicas de conteo son aquellas que sirven para enumerar eventos difíciles de cuantificar.

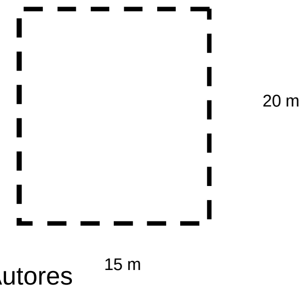
Ejemplo 3.13:

Establecer lotes de producción Cuando se establecen muestras o grupos de productos a fabricar resulta más fácil realizar el conteo unidad por unidad de los productos fabricados. Si un operario sabe que se produjeron 4 lotes en un turno de 8 horas y cada lote está formado por 12 productos, entonces se llega a la conclusión que se hizo un total de 48 productos en el turno.

Ejemplo 3.14:

Las medidas de un terreno Si se conoce la cantidad en metros de la longitud del terreno y la cantidad en metros de anchura del terreno, es posible conocer la extensión del terreno el cual equivale a la multiplicación de la longitud por el ancho del terreno.

Figura 3.15. Extensión del terreno.



Medidas del terreno (área) = 15 • × 20 • = 300 • •

Fuente: Los Autores

Las técnicas de conteo se basan en las operaciones de la suma y la multiplicación.

Entre las técnicas de conteo tenemos:

- Combinaciones
- Permutaciones
- Diagrama de árbol

3.6.1. COMBINACIONES

Una combinación es un arreglo o grupo de • elementos formados a partir de • elementos donde el orden de los mismos NO es importante, lo que se busca con una combinación es la formación de grupos y el contenido de los mismos. La combinación se calcula a través de la fórmula: $C_m^r = \frac{r!}{(r-m)! \cdot m!}$

Donde:

• > •

C_m^r = Combinaciones de • elementos tomados de entre • objetos

- Combinaciones con repetición:

Con formato: Punto de tabulación: 4 cm, Izquierda

$$CR_m^n = \frac{(m+n-1)!}{n! \cdot (m-1)!}$$

Es el caso de común juego del bingo salga 39, 40, 3, 10, 22, 49, que salga 40, 10, 39, 3, 22, 49. Además es sin repetición puesto que no puede salir dos veces el mismo número. La combinación premiada básica (6 aciertos) tiene C_6^{49} posibilidades diferentes.

3.6.2. PERMUTACIONES

Teniendo un conjunto finito con sus elementos se le llama permutación a todos los posibles casos como pueden organizar los elementos de un conjunto. Matemáticamente; •. = •!

Donde

•j: Es la permutación

•: • Elementos tomados de • en •.

!: Es la representación matemática de la operación factorial.

- **Con repetición:**
  - **Variaciones con repetición**: • elementos en grupos de •

    •• : = • •

  - **Permutaciones con repetición(pr):** ordenaciones con elementos repetidos

    •• :•,...,• • = •! / (•.! ... •.!)

Ejemplo 3.15:

Si tenemos • manzanas y deseamos colocarlas en una cesta, tendremos:

a) Si caben todas •n

b) Si no caben más de • manzanas, habría V_m^n formas diferentes de introducirlas en la cesta

3.6.3. DIAGRAMA DE ÁRBOL

Por medio de esta técnica se muestra de forma grafica la secuencia de los sucesos que se pueden realizar.

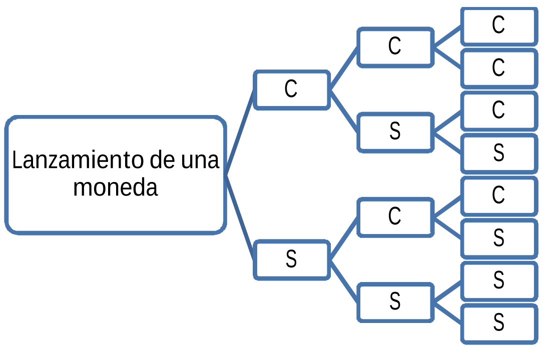
Tabla 2.2 Asignación de probabilidad.

ASIGNACIÓN DE PROBABILIDAD			
Suceso	Evento(s)	# de Combinación(es)	Probabilidad(es)
a) Lanzar una moneda	C	(c)	½ = 0.5
			½ = 0.5
	S	(s)	1 1,00
b) Lanzamiento de 2 monedas	2c	(cc)	½ = 0.25
	1c y 1s	(c) (s)	$^{2}/_{4}$ = 0.5
	2s	(ss)	½ = 0.5
			1 1,00
c) Lanzamiento de 3 monedas	3c	(ccc)	$^{1}/_{8}$ = 0.125
	2c y 1s	(cc) (s)	$^{3}/_{8}$ = 0.5
	2s y 1c	(ss) (c)	$^{3}/_{8}$ = 0.5
	2s	(sss)	$^{1}/_{8}$ = 0.125
			1 1,00

Fuente: MARTINEZ B. Ciro. Estadística y Muestreo. Doceava Edición. Bogotá DC. ECOE 2005.

La elaboración de este espacio muestral se realiza por medio de un diagrama de árbol de la siguiente manera:

Figura 3.16. Diagrama de árbol.



Fuente: MARTINEZ B. Ciro. Estadística y Muestreo. Doceava Edición. Bogotá DC. ECOE 2005.

3.7. PROBABILIDAD CONDICIONAL

Un condicional es un estado en el cual se tiene un evento en particular, dado que otro evento ya ha ocurrido o es seguro que ocurra.

Ejemplo 3.16:

- Si la temperatura aumenta, el hilo se derrite
- No habría sentido remordimiento, si te hubiera dicho la verdad

Se puede definir probabilidad condicional como “la probabilidad de que un suceso • ocurra cuando se sabe que ya ocurrió un suceso •” como:

$$P_B(A) = P(A|B) = \frac{P(B \cap A)}{P(B)}, \qquad P(B) \text{ f } 0$$

Ejemplo 3.17:

Cuando se lanzan unos dados la probabilidad de que salga cierto número es de $1/6$, si se dispone de la información anterior como que ha salido un número par en el evento anterior la probabilidad de que ahora salga un numero escogido que sea par es de $1/3$.

Se puede decir que cuando se posee información, el espacio muestral • de los posibles resultados pasa de ser • = $\{1, 2, 3, 4, 5, 6\}$ a ser • = $\{2, 4, 6\}$ y con este se trabaja.

Ejemplo 3.18:

En una planta existen dos líneas de ensamble diferentes B y B', se ensamblan componentes complejos, La línea B es de tipo manual, y la línea B' tiene un sistema automatizado, en día cualquiera en la línea B se ensamblan 17 componentes, donde se obtienen 7 componentes defectuosos (C), y en la línea B' se realizan 20 componentes, de los cuales 2 salen defectuosos (C).

Sin saber esta información, el jefe de producción elige al azar uno de los 57 componentes para realizar una demostración:

Antes de la demostración se escoge el componente de la línea:

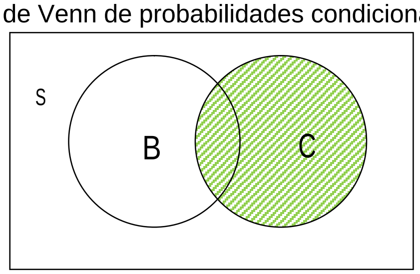
P (Elegir un componente de la línea) = $P(B)$ = $\frac{\bullet(\bullet)}{\bullet} = \frac{\bullet\bullet}{\bullet\bullet}$ = 0.45.

Si resulta defectuoso el componente elegido, entonces ocurrió el suceso C, así que el componente se encuentra entre los 9 componentes defectuosos. Puesto que estos tres componentes son equiprobables entres si después que ocurrió:

$$\bullet(\bullet|\bullet) = \frac{7}{9} = \frac{\frac{7}{37}}{\frac{2}{37}} = \frac{\bullet(\bullet\ \bullet\ \bullet)}{\bullet(\bullet)}$$

La probabilidad condicional es expresada como una relación de probabilidad incondicional, el numerador es la probabilidad de la intercepción de los dos sucesos, mientras que el denominador es la probabilidad condicional C. Esta relación es ilustrada en la siguiente figura:

Figura 3.18. Diagrama de Venn de probabilidades condicional.



Fuente: Los Autores

Dado que ocurrió C, el espacio muestral ya no es •, sino que consiste en el resultado de •; • ocurre si y solo si ocurre uno de los resultados de la intersección, así que la probabilidad condicional de B dada C es proporcional a •(• • •). La constante de proporcional 1/•(•) se usa para asegurar que la probabilidad •(•|•) del nuevo espacio muestral • es igual a 1

Ejemplo 3.19.

En el proceso productivo de la leche se pueden obtener derivados diferentes como queso y suero. La probabilidad de que el lunes se haga queso es de 0,6 y la que el martes se haga queso es de 0,3, la probabilidad de que hagan queso el lunes y el martes es de 0,1. Suponiendo que el día lunes se hace queso, ¿Cuál es la probabilidad de que se haga el martes?

Sean los eventos A: hacer queso el lunes y el evento B: hacer queso el martes. Entonces,

$$• (• | •) = \frac{• (• \ • \ •)}{• (•)} = \frac{0.1}{0.6} = 0.166$$

Se desea saber si el evento • se realiza. Cual sería la probabilidad de que el evento B no se haga:

$$1 - P(\mathrm{A} \mid \mathrm{B}) = 1 - 0{,}166 = 0{,}833$$

3.8. INDEPENDENCIA DE SUCESOS

Se dice que dos o más sucesos son independientes entre sí, si se tiene conocimiento de que cuando ocurre uno o múltiples de ellos no existe modificación alguna con la probabilidad de que ocurran los otros. Por lo cual se cumple que:

Si • y • son dos sucesos, la idea intuitiva de que son independientes seria que:

$$P (• | •) = P (•) \qquad P(• / •) = •.$$

Por lo anterior también se cumple lo siguiente:

$$P(\mathrm{B} \mid \mathrm{A}) = P(B) = \frac{P(A \cap B)}{P(A)} \quad \Rightarrow \quad P(• \cap •)=P(•)\cdot P(•)$$

Para generalizar, • sucesos son independientes, podemos decir que:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n)=P(A_1)\cdot P(A_2)\cdot \ldots \cdot P(A_n)$$

En este se cumple la independencia entre los mismos sucesos tomados de dos en dos, de tres en tres, etc.

Figura 3.19. Sucesos independientes.

Fuente: Los Autores

Evento A: Salir a caminar

Evento B: Que llueva en el día de hoy

Ejemplo 3.20

El proceso de elaboración de láminas de Dry Wall se realiza mediante 4 subprocesos, para fabricar una lámina de Dry Wall que cumpla con las especificaciones es necesario que los estándares de calidad no estén por debajo del 0,98.

Si se asume que los estándares de calidad de cada subproceso es independiente de las demás.

Se $A_{i:}$ la parte i es que se cumpla con el estándar de calidad

P (la probabilidad de obtener una lamina de Dry Wall que cumpla con las especificaciones) = $P(A_1). P(A_2). P(A_3). P(A_4)$

$P= (0,98)\ (0,98)\ (0,98)(0,98) = (0,98)^4= 0,92236816$

Ejemplo 3.21.

Un edificio cuenta con dos ascensores para el transporte de personal, debido a la demanda y a la posibilidad de falla, la probabilidad que un ascensor esté disponible es de 0,8. La disponibilidad de los ascensores es independiente una de la otra, en caso de necesitar los ascensores:

- ¿Cuál es la probabilidad de que los dos estén disponibles?
- ¿Cuál es la probabilidad que ninguno esté disponible?

Sean los eventos:
C_1: el ascensor 1 esté disponible
C_2: el ascensor 2 esté disponible

Probabilidad que estén disponibles= $P(C_1). P(C_2).= (0,8). (0,8)= 0,64$
Probabilidad que no estén disponibles= $(1- P(C_1).).(1- P(C_2)) = (0,2)\ (0,2) = 0,04$

3.9. TEOREMA DE LA INTERSECCIÓN

Sabemos que

$$P(B|A)=\frac{P(A\cap B)}{P(A)} \qquad P(A|B)=\frac{P(A\cap B)}{P(B)}$$

Si se despeja la intercepción de cada una de las ecuaciones se obtiene lo siguiente:

P (A∩B)=P(A|•)·P(B)=P(B|•)·P(A)

Este es llamado *Teorema de la Intersección* para dos sucesos. El siguiente es para tres sucesos, se puede colocar la condicional como:

$$P(C/B/A)=\frac{P(C\cap B/A)}{P(B/A)}=\frac{P(A\cap B\cap C)}{P(A\cap B)}=P(C/(A\cap B))$$

El teorema de intercepción tres sucesos es:

$$P\ (A\cap B\cap C) = P\ (B)\ P\ (A/B)\ P\ (C/A\cap B)$$

En generalidad para un número de • sucesos es el siguiente:

Sean •., •., ... , •. n sucesos tal que •(•.|•. • ... • •...) > 0,entonces:

•(•. • ... • •...) = •(•.)•(•.|•.)•(•.|•. • •.) ... •(•.|•. • ... • •...)

Ejemplo 3.22

En una urna se tienen 4 esferas azules y 2 esferas rojas. Si se extraen dos esferas al azar de la urna. Cuál es la probabilidad de que:

Ambas sean azules

La primera sea rojas y la segunda azul

Todas las extracciones se hacen sin remplazo. Los eventos son los siguientes:

•1: Obtener una esfera azul en la primera extracción

•2: Obtener una esfera azul en la segunda extracción

•3: Obtener una esfera roja en la primera extracción

•4: Obtener una esfera roja en la segunda extracción

P (las 2 esferas azules) = $P(A_\bullet \bullet A_\bullet) = P(A_\bullet) \times \bullet\frac{\bullet\bullet}{\bullet\bullet}\bullet = \frac{\bullet}{\bullet} \times \frac{\bullet}{\bullet}$ = 0.4

P (La primera esfera roja y la segunda azul) = $P(A_\bullet \bullet A_\bullet)$

$= P(A_\bullet \times \; P(A_\bullet / A_\bullet)) = \frac{\bullet}{\bullet} \times \frac{\bullet}{\bullet}$ = 0.2666

Si las extracciones se realizan con reemplazo:

P (las 2 esferas azules) = $P(A_\bullet \bullet A_\bullet) = P(A_\bullet) \times \bullet\frac{\bullet\bullet}{\bullet\bullet}\bullet = \frac{\bullet}{\bullet} \times \frac{\bullet}{\bullet}$ = 0.444

P (la primera esfera roja y la segunda azul) = $P(A_\bullet \bullet A_\bullet) = P(A_\bullet \times \; P(A_\bullet / A_\bullet) =$ $\frac{\bullet}{\bullet} \times \frac{\bullet}{\bullet}$ = 0.2222

Ejemplo 3.23

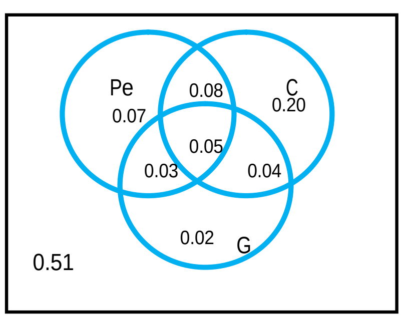
La refinería de Cartagena produce tres tipos de productos Petróleo (Pe), Crudo (C), Gasolina (G), se escoge el porcentaje de ventas de un semestre con relación a los productos descritos anteriormente

Tabla 3.3 Probabilidad de Comprar

Pe	**C**	**G**	**Pe•C**	**Pe•G**	**G•C**	**Pe•G•C**
0.23	0.37	0.14	0.13	0.08	0.09	0.05

Fuente: Los Autores

Figura 3.21 Diagrama de Venn del porcentaje de ventas.



Fuente: Los Autores

Por lo tanto se tiene que:

$$\bullet(\bullet|\bullet\bullet) = \frac{\bullet(\bullet\ \bullet\ \bullet\bullet)}{\bullet(\bullet)} = \frac{0.08}{0.23} = 0.35$$

$$\bullet(\bullet|\bullet\bullet\ \bullet\ \bullet) = \frac{\bullet(\bullet\ \bullet\ (\bullet\bullet\ \bullet\ \bullet)}{\bullet(\bullet\bullet\ \bullet\ \bullet)} = \frac{0.04 + 0.05 + 0.03}{0.47} = \frac{0.12}{0.47} = 0.26$$

$$\bullet(\bullet|\bullet\bullet\ \bullet\ \bullet\bullet\bullet\ \bullet\bullet\bullet\ \bullet\bullet\ \ \bullet\bullet\bullet\bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet\bullet\bullet) = \bullet(\bullet|\bullet\ \bullet\ \bullet\bullet\ \bullet\ \bullet) = \frac{\bullet(\bullet\ \bullet\ (\bullet\ \bullet\ \bullet\bullet\ \bullet\ \bullet)}{\bullet(\bullet\ \bullet\ \bullet\bullet\ \bullet\ \bullet)}$$

$$= \frac{\bullet(\bullet)}{\bullet(\bullet\ \bullet\ \bullet\bullet\ \bullet\ \bullet)} = \frac{0.14}{0.49} = 0.29$$

3.10. PROBABILIDAD DE INTERSECCIÓN DE SUCESOS INDEPENDIENTES:

Si • y • son dos sucesos independientes, se verifica que la probabilidad de la intersección de A y B es igual al producto de las probabilidades de cada uno de los sucesos:

$$\bullet(\bullet\ \bullet\ \bullet) = \bullet(\bullet)\bullet\bullet(\bullet)\ \ \bullet\bullet\ \bullet\ \bullet\ \bullet\ \bullet\bullet\bullet$$

Ejemplo 3.24:

En una empresa productora de relojes se tiene un lote de 40 productos, los cuales se encuentran distribuidos de la siguiente forma: 20 relojes grandes, 15 relojes pequeños, y los demás medianos, se extrae uno y se vuelve a meter. ¿Cuál es la probabilidad de extraer dos medianos?

$$\bullet(\bullet_{.}\ \bullet\ \bullet_{.}) = \bullet(\bullet_{.})\bullet\bullet(\bullet_{.}) = \frac{5}{40}\bullet\frac{5}{40} = \frac{5}{320}$$

3.11. PROBABILIDAD DE INTERSECCION DE SUCESOS DEPENDIENTES:

Es la probabilidad de que uno de los sucesos, multiplicado por la probabilidad del segundo sabiendo que ha ocurrido el primero

:

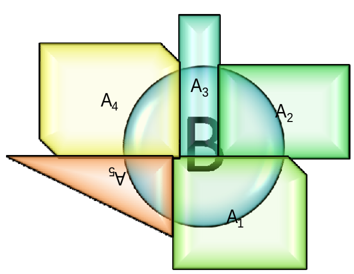
Ejemplo 3.25

Se tiene el mismo lote del ejemplo anterior pero ahora, se extraen dos relojes. ¿Cuál es la probabilidad de extraer dos ases?

3.12. TEOREMA DE LA PARTICIÓN O DE LA PROBABILIDAD TOTAL.

Se tiene la partición de un espacio muestral, es decir, n sucesos que cumplen que $\cup A_i=E$ y $A_i \cap A_j=\varnothing$ $\forall$ $i \neq j$, como se muestra en la siguiente figura:

Figura 3.22. Teorema de la partición.



Fuente: Los Autores

Donde el suceso B es la unión de:

$B = \cup(A_i \cap B)$ con $(A_i \cap B) \cap (A_j \cap B) = \emptyset \quad \forall\ i \neq j$

Según el tercer axioma de la probabilidad, podemos hallar la probabilidad de que ocurra el suceso B como:

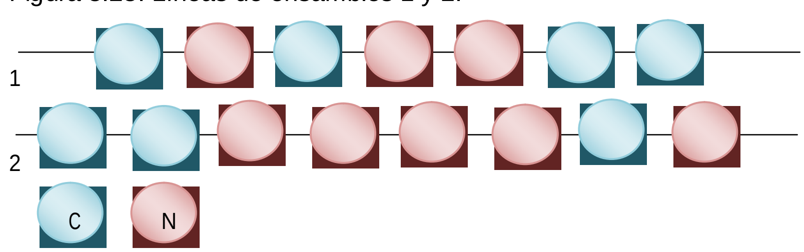
$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$$

Este es llamado como Regla de la Probabilidad Total.

Ejemplo 3.26:

Consideremos dos líneas de ensambles de productos, con productos no conformes y productos conformes. La primera contiene 3 no conformes y 4 conformes, la segunda 5 no conformes y 3 conformes. Sacamos un producto de la línea número 1 y sin ver su estado lo introducimos en la línea número 2 y a continuación extraemos de ésta un producto. Nos preguntamos por la probabilidad de que este segundo producto sea no conforme.

Figura 3.23. Líneas de ensambles 1 y 2.

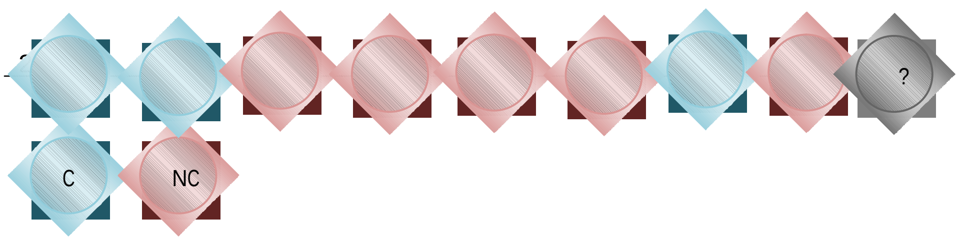


Fuente: Los Autores

Primeros denotaremos los productos como “C” es un producto conforme y “NC” producto No Conforme, acompañado de los subíndices para indicar la línea de ensamble ya sea 1 ó 2.

Entonces tenemos que NC_1 es el suceso de sacar el primer producto no conforme de la línea No. 1, y con NC_2 se quiere saber el segundo suceso que el primer producto sale no conforme, pero se tiene en cuenta que la composición de la segunda línea de ensamble cambia en el momento que le introducen el producto

Figura 3.24. Líneas de ensambles 1.



Fuente: Los Autores

Por lo cual tendríamos un total de 9 productos y el número de productos conformes y no conformes dependen del extraído de la línea #1, por lo tanto la probabilidad del suceso NC_2 depende si en el primer suceso se extrajo NC_1 ó C_1, pero los dos sucesos no pueden acontecer al mismo tiempo.

$$\bullet(\bullet\bullet_{.}) = \bullet(\bullet\bullet_{.})\bullet\bullet(\bullet\bullet_{.}|\bullet\bullet_{.}) + \bullet(\bullet_{.})\bullet\bullet(\bullet\bullet_{.}|\bullet_{.}) =$$

$$\frac{3}{7}\bullet\frac{6}{9}+\frac{4}{7}\bullet\frac{5}{9}=\frac{38}{63}$$

3.13. TEOREMA DE BAYES[8]

[8] *Tomado y Adaptado de Curso De Especialista En Gestión Y Control De La Calidad Departamento De Estadística E Investigación Operativa*, *Universidad Politécnica de Valencia, Autores Andrés Carrión García Mª Teresa Carot Sánchez*

El teorema de Bayes es el que determina la probabilidad de que la causa A_i haya sido provocada por un efecto B, en donde A_i son las llamadas causas de un suceso B que se desempeña como el efecto producido.

Si los A_i son una partición del espacio muestral y el suceso B es de probabilidad no nula, se cumple que:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)}$$

Aplicando los expuestos anteriormente los cuales son Teorema de la Intersección en el numerador y el Teorema de la Partición en el denominador, obtenemos que:

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{i=1}^{n} P(A_i) \cdot P(B|A_i)}$$

3.14 DETERMINACION DE PROBABILIDADES A TRAVÉS DE EXCEL.

Excel dispone de funciones estadísticas que permiten determinar las probabilidades de ocurrencia de un suceso o evento.

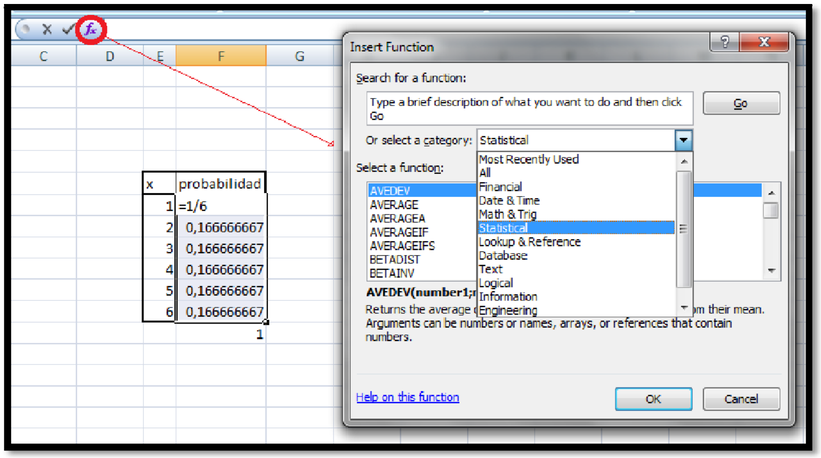
Para indicar como se calcula las probabilidades en Excel se realizará el siguiente ejemplo:

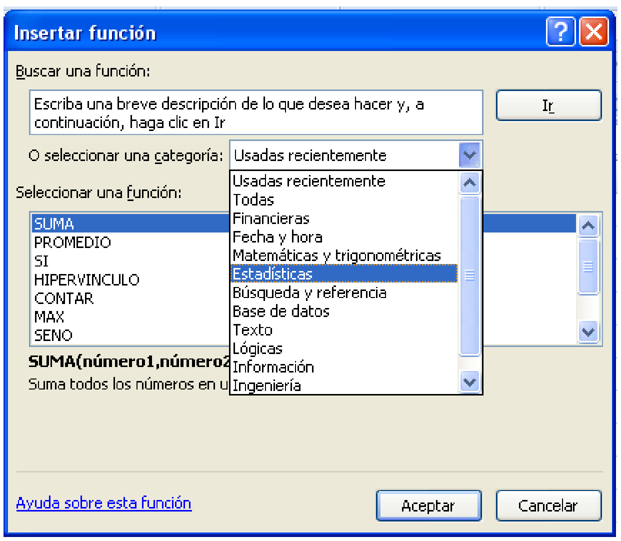
Ejemplo 1.

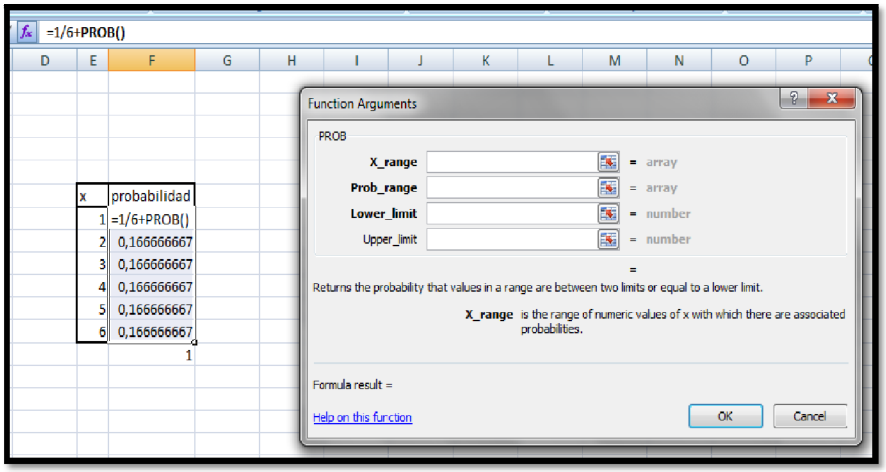
Si deseamos calcular la probabilidad de cierto evento, como es el lanzamiento de dados, primero se introducen todos los datos en la hoja de cálculo.

x	probabilidad
1	0,166666667
2	0,166666667
3	0,166666667
4	0,166666667
5	0,166666667
6	0,166666667

Después se busca en funciones la función de probabilidad, y en esta se selecciona la categoría de estadística, y dentro se busca la función de probabilidad.

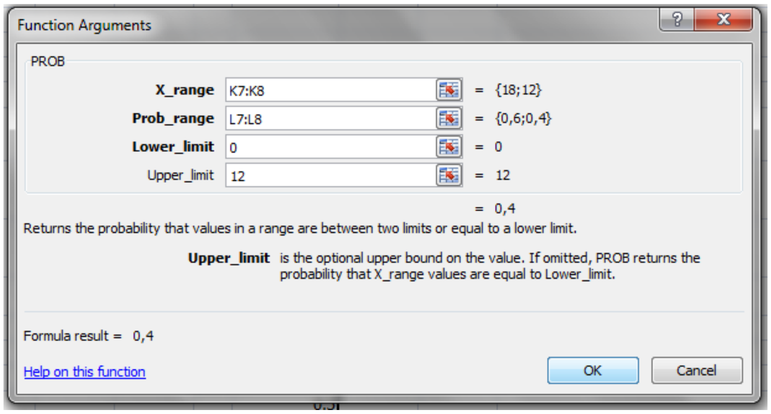

Insert Function
Search for a function:
Type a brief description of what you want to do and then click Go
Go
Or select a category: Statistical
Select a function:
Most Recently Used
All
Financial
Date & Time
Math & Trig
Statistical
Lookup & Reference
Database
Text
Logical
Information
Engineering
AVEDEV
AVERAGE
AVERAGEA
AVERAGEIF
AVERAGEIFS
BETADIST
BETAINV
AVEDEV(number1;
Arguments can be numbers or names, arrays, or references that contain numbers.
Help on this function
OK
Cancel
x
probabilidad
1
=1/6
2
0,166666667
3
0,166666667
4
0,166666667
5
0,166666667
6
0,166666667
1


Insertar función
Buscar una función:
Escriba una breve descripción de lo que desea hacer y, a continuación, haga clic en Ir
Ir
O seleccionar una categoría: Usadas recientemente
Seleccionar una función:
Usadas recientemente
Todas
Financieras
Fecha y hora
Matemáticas y trigonométricas
Estadísticas
Búsqueda y referencia
Base de datos
Texto
Lógicas
Información
Ingeniería
SUMA
PROMEDIO
SI
HIPERVINCULO
CONTAR
MAX
SENO
SUMA(número1,número2
Suma todos los números en u
Ayuda sobre esta función
Aceptar
Cancelar

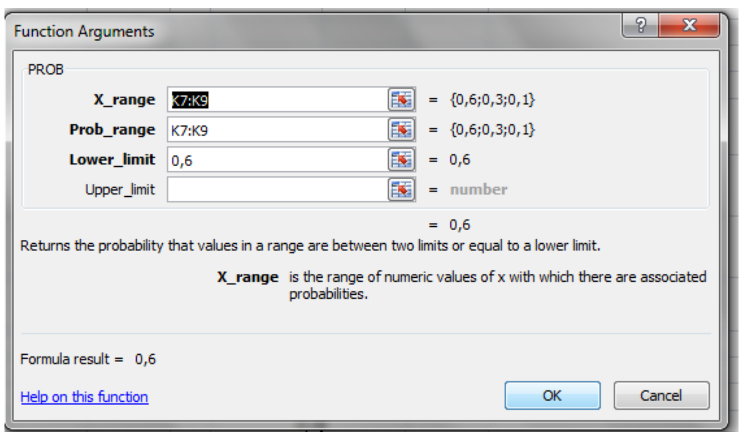


Se coloca el rango x, después el rango de la probabilidad que debe ser mayor que 0, pero siempre menor que 1, se colocan los limites de donde se quiere que llegue, el inferior y el superior. Queda de la siguiente manera; como se muestra en la siguiente figura:

Y este daría el resultado de la probabilidad que estamos hallando, tomando el segundo caso de del ejemplo No. 3.2, podemos hacerlo de manera que primeros nos de los resultados deseados y después se le reste este resultado a uno para que hallemos el deseado, si es así no se coloca el límite superior como se muestra en la siguiente figura.



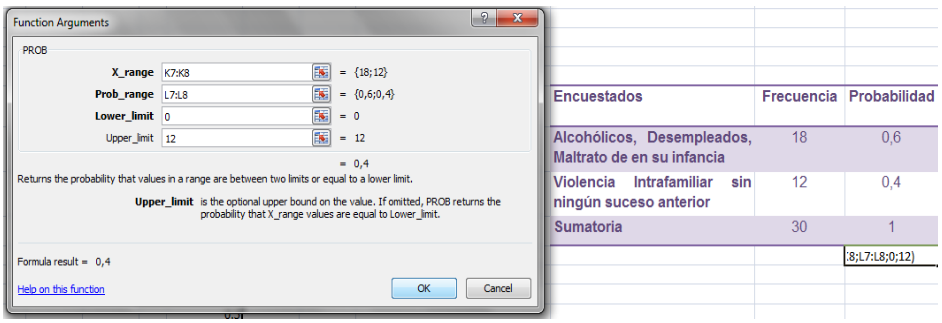
Pero si se desea obtener el resultado de manera directa se toma dándole como límite inferior 0 y como límite superior 12.



Ejemplo 2

En el proceso productivo de la leche se pueden obtener derivados diferentes como queso y suero. La probabilidad de que el lunes se haga queso es de 0,6 y la que el martes se haga queso es de 0,3, la probabilidad de que hagan queso el lunes y el martes es de 0,1. Suponiendo que el día lunes se hace queso, ¿Cuál es la probabilidad de que se haga el martes?

Se tomo la función de probabilidad para buscar la unión correspondiente, y se documentan los datos en los requisitos según lo muestra la figura.



EJERCICIOS PROPUESTOS

CAPÍTULO N° 3.

EJERCICIOS PROPUESTOS

1. Dar ejemplo de tres eventos aleatorios

2. En el lanzamiento de dos dados, se observa el número de puntos que aparecen en las caras superiores. Especifique los elementos del espacio muestral y los siguientes eventos:

   A. La suma de los puntos es igual a 7
   B. Uno de los dados muestra 4 puntos
   C. La suma de los puntos es menos a 4
   D. Los eventos A y B se dan de manera simultánea
   E. Los eventos B y C se dan de manera simultanea
   F. Se puede presentar el evento A o B

3. Del problema anterior calcular la probabilidad para los sucesos C, D, E, y F.

4. Una caja contiene 3 pares de zapatos (6 zapatos), ¿cuál es la probabilidad que en la primera extracción de la caja se saquen los zapatos que son pares?

5. En una fábrica se selecciona al azar 10 cantidades de un productos y de estos se seleccionan 2 muestras del producto para evaluar los estándares de calidad del mismo, si después de ser evaluados no se encuentra algún producto no conforme ¿Cuál es la probabilidad de que en el grupo de 10 productos se acepten 3 productos no conformes?

6. Ay B son equipos de voleibol y juegan cinco sets. Suponiendo que no hay empate, determine los posibles resultados de los tres sets teniendo en cuenta tanto al ganador como al perdedor de cada una. Ahora Suponiendo que A gana por lo menos dos partidas, ¿cuál es la probabilidad de esta?

7. Un Mago dice que puede adivinar por medio de sus poderes la presencia de petróleo en el suelo, para corroborar dicha información se han colocado seis recipientes que contienen lo siguiente: 2 agua, 2 petróleo, 1 aceite, 1 gaseosa; para que el mago determine cuál es el que contiene petróleo; si los poderes del mago son mentira, Cual es la probabilidad de que el mago identifique correctamente los dos recipientes que contienen petróleo?

8. Se acerca septiembre y seis amigos deciden jugar al amigo secreto, para esto colocan sus nombres en un papel y los introducen dentro de una bolsa; Cada persona selecciona un papel al azar, cual es la probabilidad de que a) Uno de los amigos le salga su nombre, b) dos de los amigos elijan sus propios nombres.

9. En un grupo de 15 personas se encuentran 5 menores de edad, y estos desean entrar a una discoteca (Lo cual no se permite por ser menores de 18 años y va contra la ley). De repente cuando se acercan a la portería de la discoteca el portero se les acerca. A) calcule la probabilidad de que el portero no deje entrar a uno de los menores, b) a 3 de los menores, c) a los 5 menores

10. 4 mujeres y 4 hombres se acomodan en una fila de 8. Si eligen sus puestos al azar cual es la probabilidad : a) de que queden de manera intercalada, b) de que los hombres queden juntos, c) de que queden 2 hombres 2 mujeres 2 hombres 2 mujeres.

11. Calcule la probabilidad de obtener 7 cuando se lanzan los dados tres veces?

12. En un experimento de sicología no se pueden utilizar individuos que sean zurdos, en la población en donde se va a aplicar dicho experimento cada 7 personas 2 son zurdas. Si se escoge al azar un individuo, ¿Que probabilidad existe que no pueda participar en dicho experimento?

13. En una urna se introducen 7 esferas blancas, 6 esferas rojas, 3 esferas negras, 3 esferas naranja. Si se extraen 3 esferas al azar al mismo tiempo de la urna. ¿Calcular la probabilidad de que estas sean del mismo color?

14. En un estudio de control estadístico de procesos se realizaron experimentos, en donde se extrajeron muestras de un liquido cada 3 horas durante una semana (7 días), existe la probabilidad que el 35% de las muestras se encuentren defectuosas ¿Calcule la probabilidad de que en 15 días se extraigan 2 muestras que no se encuentren defectuosas?

15. El jefe de control de calidad y de control de procesos revisan un artículo para catalogarlo como bueno o defectuoso. El primero de los jefes no detecta como defectuosos el 7% de los artículos revisados. Y el segundo jefe no detecta el 2% de los artículos defectuosos detectados por el primero de los jefes. Calcule la probabilidad de que un articulo defectuoso pueda pasar por los estándares de calidad sin ser detectado?

16. Se realizo un estudio entre un grupo de personas con relación a las alzas en los diferentes servicios públicos (todo eso para sopesar los impuestos para la guerra), el 18% de las personas que interfirieron en el estudio estuvieron de acuerdo con el alza, y de eso el 55% eran personas las cuales sus trabajos se encontraban relacionados con la política. Calcule la

probabilidad de que una persona que se encuentre trabajando en el área de la política este de acuerdo con el alza?

17. Dos personas se encuentran de caza y visualizan un pato, estos disparan al mismo tiempo disparan. La probabilidad de que el primer tirador le de al pato es 4/5, y que el segundo tirador le de al pato es de 3/4. Calcule la probabilidad de que el animal quede muerto?

18. Según el Ministro Empleo en Colombia el 30% de la población tiene empleos de tipo informal, el 14% de la población se encuentra desempleada, si se eligen 7 personas al azar. Calcular la probabilidad de que los 7 se encuentre desempleados, que 3 tengan empleo de tipo informal y los restantes se encuentren desempleados, 2 de ellos tenga trabajo de tipo informal, y por último que ninguno se encuentre entre estos porcentajes?

19. Una persona tiene que llegar al odontólogo y al supermercado a tiempo. La probabilidad de llegar al odontólogo a tiempo es de 0.8% y de llegar a tiempo al supermercado es de 0.1, Sabiendo que se retraso al momento de llegar al odontólogo, cual es la probabilidad de que a pesar del retraso se llegue a tiempo al supermercado?

ESTUDIO DE CASOS

CASO 1

En la empresa “CONCRETOS CARTAGENA” dedicada al suministro de concreto se realizan los pedidos con un día de anterioridad, y en la mañana con un intervalo de 7 a.m. hasta las 9 a.m., este tipo de pedidos que entran en el horario diurno no pueden estar entre los pedidos diurnos (7a.m a 12 a.m) a menos que algunos de clientes que tienen pedidos diurnos lo cancele, la probabilidad de que un pedido sea cancelado es de 0.25, y la probabilidad de que entre algún pedido solicitado en horario de 7 a.m. a 9 a.m. entre es de 0.15; en la empresa se despachan aproximadamente 16 viajes de concreto, y cuentan con 5 carros para el suministro de concreto, la capacidad de los carros es de 7Mts3.

En la empresa existe un problema con los tiempos de entrega del producto, ya que en muchas ocasiones influye en la entrega el estado del tiempo, la probabilidad de que llueva en un mes es de 0.066 en las épocas calurosas y en las épocas de lluvia es de 0.26, además que es necesario tener en cuenta los siguientes puntos:

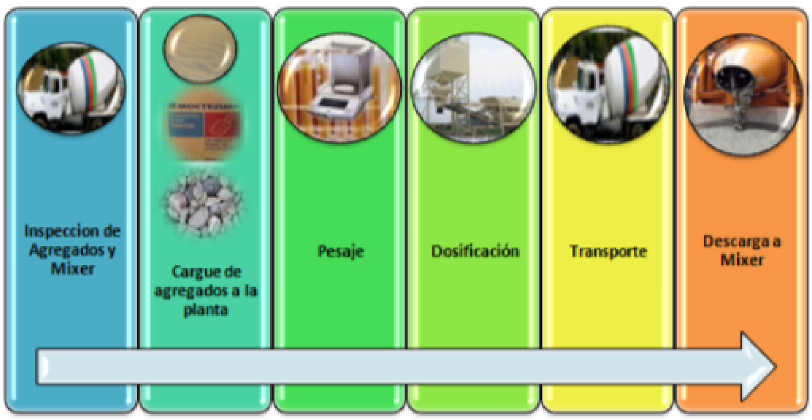
Los tiempos de alistamiento (TA): es el tiempo que le toma al operario realizar la acción de colocar debajo de la planta que suministra el concreto el vehículo para iniciar el proceso de cargue, inspección de agregados, y cargue de agregados a la planta dosificadora (Min)
Los tiempos de producción (TP): tiempo que toma realizar la mezcla de grava, arena, cemento, agua y aditivos para obtener el concreto (min)
Tipos de concreto los cuales tienen tiempos de descargue.
Velocidad recorrida: 12 Km/hr, esta velocidad no es variable

Tabla 3.4 Datos LMTM

Tipo de Concreto	Tiempo de Alistamiento Min	Tiempo de producción min/mts3	Tiempo de descargue min/mts3
2000 PSI	2.52	1.12	10.71
3000 PSI	2.15	1.01	29.55
3500 PSI	2.11	1.59	38.12
MR – 39	1.15	0.47	8.54
MR – 45	1.02	0.52	5.27

Fuente: Datos tomados por un estudio de métodos y tiempos de la empresa LMTM

Figura 3.25 Diagrama de Proceso de la Producción de Concreto



Fuente: Datos tomados por un estudio de métodos y tiempos de la empresa LMTM

Lo siguiente es una programación:

Horario	**Tipo de Concreto**	**Cantidad (• ˙)**	**Distancia de Recorrido (Km)**
7:00	3000 PSI	5	5.2
7:30	2000 PSI	7	3
8:00	MR-45	7	2
8:30	MR-45	4	8
9:00	MR-45	4	8
9:30			
10:00	MR-39	3.5	5.6
10:30	3000 PSI	6	10
11:00	2500 PSI	4.2	1
11:30			
14:00	3000 PSI	2	3.27
14:30	MR-39	9	15
15:00	MR-45	7	3.5
15:30	3000 PSI	1	10
16:00	3000 PSI	9	3
16:30	3500 PSI	12	13

Fuente: Datos tomados por un estudio de métodos y tiempos de la empresa LMTM

Teniendo en cuenta que la época del año es soleada:

- Realiza un estudio con los datos anteriores determinando las probabilidad de puntualidad de los pedidos si existe o no, y plante una nueva programación para este tipo de pedidos mostrando una posible solución a este problema, en el caso que no se pueda dar una solución que propuesta de mejora se podría aplicar a la situación, además, plante con algunos de los tipos de concreto un despacho continuo simultaneo en donde estos lleguen a tiempo, recuerde este tipo de estudio se realiza para épocas soleadas y lluviosas

- En el caso que ha sido cancelado un pedido para las 9:30, ¿Cuál es la probabilidad que un pedido de 3500 PSI de 2.5 Mts3m que recorra una distancia de 5 km, llegue a tiempo? ¿Existe alguna probabilidad para determinar que se cuenta con el número de carros para despachar este pedido? De lo contrario, justifique su respuesta.

CASO 2

Siguiendo con el Caso No. 2 del Capítulo 2, tenemos el siguiente formulario Utilizado en Cartagena Como Vamos.:

FORMULARIO DE OBSERVACIONES

CARTAGENA COMO VAMOS

<table>
<tr><td colspan="4">Plantilla de Movilidad Vial
FECHA: ______________</td></tr>
<tr><td>Trayecto*</td><td colspan="3"></td></tr>
<tr><td>Lugar de Salida</td><td></td><td>Hora de Salida</td><td></td></tr>
<tr><td colspan="2">Ruta por Tramos</td><td colspan="2">Tiempo registrado (HH-MM)</td></tr>
<tr><td>Lugar de Llegada</td><td></td><td>Hora de llegada</td><td></td></tr>
<tr><td>Duración total del Trayecto</td><td colspan="3"></td></tr>
<tr><td>Observaciones</td><td colspan="3"></td></tr>
<tr><td colspan="2">1. Cuantos pasajeros hay en el bus al momento en que usted lo aborda</td><td colspan="2"></td></tr>
<tr><td colspan="2">2. El vehículo tiene los asientos completos</td><td colspan="2">a) Si
b) No</td></tr>
<tr><td colspan="2">3. El volumen que lleva el radio del vehículo es</td><td colspan="2">a) Bajo
b) Alto
N.A. – No tiene el radio encendido</td></tr>
<tr><td colspan="2">4. El vehículo dispone de timbre para solicitar la parada</td><td colspan="2">a) Si a1) Funciona a2) No funciona
b) No</td></tr>
<tr><td colspan="2">5. Hay extinguidores a la vista</td><td colspan="2">a) Si
b) No</td></tr>
<tr><td colspan="2">6. Tiene el vehículo salidas de emergencia señalizadas</td><td colspan="2">a) Si
b) No</td></tr>
<tr><td colspan="2">6ª Junto a las salidas de emergencia hay instrucciones indicando la forma en que deben ser usadas</td><td colspan="2">a) Si
b) No</td></tr>
<tr><td colspan="2">7. Durante el recorrido, las puertas de vehículo estuvieron:</td><td colspan="2">a) Abiertas
b) Cerradas</td></tr>
<tr><td colspan="2">8. Presentó el vehículo alguna falla mecánica que retrasara el recorrido</td><td colspan="2">a) Si Debió: a) Hacer trasbordo
B) No b) Esperar para continuar en el mismo bus</td></tr>
<tr><td colspan="2">9. Independiente de una falla mecánica, ¿debo hacer trasbordo de vehículo?</td><td colspan="2">a) Si
b) No</td></tr>
<tr><td colspan="2">Código encuestador</td><td colspan="2"></td></tr>
<tr><td colspan="2">Coordinador a cargo</td><td colspan="2"></td></tr>
</table>

A continuación se muestra el resultado para el año 2009 de esta encuetas

2. Que vehículos tiene los asientos completos?

   Si (244 vehículos) No (5 Vehículos)

3. El volumen que lleva el radio del vehículo es?

   Alto (66 vehículos) Bajo (115 vehículos) NA (36 vehículos)

4.a. El vehículo dispone de timbre para solicitar la parada?

   Si (195 vehículos) No (52 vehículos)

4.b. El vehículo del timbre…

   Funciona (114 vehículos) No Funciona (47 vehículos)

5. Hay extintores a la vista

   Si (28 vehículos) No (219 Vehículos)

6.a. Tiene el vehículo salidas de emergencia señalizadas

   Si (201 vehículos) No (25 Vehículos)

6.b. Junto a las salidas de emergencia hay instrucciones indicando la forma en que deben ser usadas.

   Si (124 vehículos) No (123 Vehículos)

7. Durante el recorrido, las puertas del vehículo estuvieron

   Si (243 vehículos) No (5 Vehículos)

8.a Presentó el vehículo alguna falla mecánica que retrasara el recorrido

   Si (5 vehículos) No (243 Vehículos)

8.b. Debió hacer transbordo por la falla mecánica

   Si (5 vehículos) No (0 Vehículos)

9. Independiente de una falla mecánica. Debió hacer trasbordo del vehículo?

   Si (2 vehículos) No (3 Vehículos)

De acuerdo a los resultados:

1. Relaciona cada uno de los aspectos abordados en el estudio y determina la probabilidad que estos aspectos se cumplan o no.
2. Realice un balance general del estado de los vehículos utilizados para el transporte público en Cartagena, tenga en cuenta la seguridad que proporcionan, la comodidad y demás aspectos abordados en las preguntas.

CAPÍTULO 4

VARIABLES ALEATORIAS

INTRODUCCION

El concepto de variables aleatoria constituye uno de los pilares básicos de la teoría de la probabilidad.

El concepto de variable aleatoria estuvo presente desde el inicio de la teoría de la probabilidad. Huygens introdujo una variable aleatoria que sigue una distribución Hipergeometrica para resolver uno de los problemas de su libro. Cuando Galileo habló de los errores ‘aleatorios’ que no se pueden predecir y que varían de medida en medida, en el cual los errores constituyen una variable aleatoria de distribución desconocida.

Cuando Bernoulli enunció su ley de los grandes números, al contabilizar el número bolas blancas extraídas de la urna, ese número de éxitos es una variable aleatoria que toma valores entre 1 y n el número total de pruebas, siguiendo una distribución binomial.

En el comienzo del siglo XIX, se utilizaba ampliamente la distribución normal para resolver problemas que surgían. Sin embargo autores como De Moivre, Gauss, Daniel Bernoulli o Laplace mencionaron el concepto de variables aleatorias.

El primero en mencionar la idea de variable aleatoria fue Poisson en el año de 1832, en su libro sobre la Probabilidad de los Resultados Promedios de Observaciones. No utilizó el término ‘variable aleatoria’, pero introdujo la idea de las variables aleatorias discretas al considerar la relación que existe entre un conjunto a_1, a_2, …, a_n con sus correspondientes probabilidades p_1, p_2, …, p_n. También, consideró variables aleatorias continuas y sus densidades.

La expresión ‘variable’ fue utilizada por primera vez por Chebyshev, que asumió implícitamente que todas las variables aleatorias eran independientes y fue A.

Liapunov (1857–1918) el primero que usó sistemáticamente el término ‘variable aleatoria’ y especificó que serían independientes cuando fuese necesario.

En el comienzo de su obra sobre una afirmación en la teoría de probabilidad, Liapunov dio la definición de función de distribución tal y como la conocemos hoy: •(• < • < •) = • (•) – • (•).

El estudio del comportamiento de una variable aleatoria se origina de su medición cuantitativa obteniendo valores que conforman un espacio muestral que resulta de un ensayo, experimento o fenómeno aleatorio.

En muchos problemas es necesario determinar las probabilidades de que una variable aleatoria describa valores específicos dentro de un rango de valores posibles.

Por lo anterior, en este capítulo se incluye el concepto de variable aleatoria, tipos de variables aleatorias, función de probabilidades para cada tipo de variable aleatoria. Además del valor medio o esperanza, la varianza y desviación típica, que son propiedades importantes de una variable aleatoria ya que permiten estimar parámetros de la población para determinar su comportamiento.

COMPETENCIAS:

- Interpretacion e identificación de variables aleatorias discretas y continuas en un experimento aleatorio
- Calculo del valor esperado, varianza y desviación típica para variables aleatorias discretas y continuas.
- Interpretación de valores esperado, varianza y desviación típica para variables aleatorias discretas y continuas.
- Resolucion y formulacion de problemas determinando la función de distribución que describe una variable aleatoria discreta o continua
- Resolucion y planteamiento de problemas usando las diferentes distribuciones para variables aleatorias discretas, argumentando el uso de cada una de ellas.

4. VARIABLE ALEATORIA

Las variables aleatorias corresponden a una caracterización cualitativa de los resultados que constituyen un espacio muestral, cada cantidad o valor es el resultado de un experimento aleatorio, como tal, puede tomar distintos valores.

Se dice que se define una variable aleatoria para un experimento aleatorio cuando se asocia un valor numérico a cada resultado del experimento.

En este sentido, Sea • el espacio muestral asociado a un experimento, se llama variable aleatoria a toda aplicación del espacio muestral • en el conjunto de los números reales •

Para designar variables aleatorias, se utilizan letras mayúsculas •, •, • ... y las letras minúsculas •, •, • ... para designar valores de la variable.

El valor de la variable aleatoria, depende del resultado del experimento, una variable aleatoria puede ser discreta o continua, depende del tipo de valores numéricos que asuma.

Sin embargo, es importante destacar que en aquellos experimentos en los cuales se realiza conteo, el valor numérico de la variable que se obtiene es de tipo discreto. Mientras, que en los experimentos en los cuales se realiza medición, la variable es continua.

4.1. VARIABLE ALEATORIA DISCRETA

Las variables aleatorias discretas son aquellas que toman un número finito o infinito numerable de valores.

Ejemplo 4.1*:*

a) Una empresa se dedica a la producción de laminas de Drywall, fabrica al día aproximadamente 2000 placas,

   Se define la Variable aleatoria discreta como *X*= *"No. De laminas defectuosas"*
   X puede tomar valores desde 0, 1, 2, 3… 2000.
   Esta variable en la manera como se encuentra definida es de tipo discreta.

b) En una bolsa se encuentran 10 canicas de color rojo y 10 canicas de color azul. Se extrae al azar una canica. No se sabe con certeza el color de la canica que se extrajo, este experimento se sigue realizando 6 veces más.

   Al terminar de realizar el experimento se desea conocer cuantas canicas de cada color se extrajeron. Para esto Definimos las variables aleatorias:

   X = "número de canicas de color azul que se extrajeron de la bolsa"
   Y = *"número de canicas de color rojo que se extrajeron de la bolsa"*

   Las variables aleatorias así definidas son discretas siendo su valor inicial 0, decir que no se extrajo ninguna canica de determinado color y para cada una de las variables X o Y, el máximo valor que puede tomar es 7. Debido que se realizan 7 extracciones en total.

4.2. VARIABLE ALEATORIAS CONTINUAS

Las variables aleatorias continuas son aquellas que pueden obtener cualquier valor en la recta de los reales.

Los resultados de experimentos basados en escala de medición del peso, tiempo, temperatura, distancia pueden ser descritas por variables aleatorias continuas.

Es importante destacar que el estudio de las variables aleatorias continuas requiere del cálculo matemático, las integrales y las derivadas.

Ejemplo 4.2**:**

Para el proceso de fabricación de las galletas, las materias primas deben ser almacenadas a temperaturas reguladas para su refrigeración.

La refrigeración es necesaria para la conservación de la materia prima, esta se debe realizar a una temperatura entre 4º C y 8º C. Para esto el operario debe medir constantemente la temperatura.

Definamos X como la temperatura de refrigeración, esta puede asumir valores como 4,5º C; 5,8º C. en general: 4º C • X • 8º C.

Ejemplo 4.3:

En la pizzería Paco's se realiza un estudio para determinar el tiempo en espera promedio que demora un cliente en ser atendido una vez el ingresa a la cola para realizar su pedido.

Este estudio se realiza con el fin de lograr aumentar la satisfacción de los clientes, reduciendo los tiempos de espera.

Definamos Y como el tiempo de permanencia de un cliente en la cola.
Para un primer cliente C1, se cronometra y se obtiene un tiempo de 1,3 minutos en la cola.

Esto se hace con el resto de clientes que llegan en determinado rango de hora a la pizzería Paco’s.

De acuerdo al estudio el tiempo mínimo que puede permanecer un cliente en cola es de 1 minuto y el tiempo máximo es de 10 minutos. Por lo cual:

$$1 \bullet \ \bullet \bullet \ 10$$

4.3. EL VALOR ESPERADO DE UNA VARIABLE ALEATORIA

En estadística, el valor esperado también conocido como la esperanza matemática, media o esperanza de una variable aleatoria es el número que describe el valor medio en un experimento.

4.3.1. El Valor Esperado En Variables Aleatorias Discretas

Si X toma los valores $x_\bullet, x_\bullet, \ldots, x_\bullet, \ldots, x_\bullet$ con probabilidades $p_\bullet, p, \ldots, p_\bullet, \ldots, p_\bullet$, Se establecerá la esperanza o media como lo siguiente:

$$E[X] = \overset{\bullet}{\underset{\bullet\bullet\bullet}{\bullet}} x_\bullet \bullet p_\bullet$$

Propiedades:

- Una variable aleatoria que describe un valor constante C, en otras palabras la variable es constante. Su esperanza es esa misma constante.

$$\bullet[\bullet] = \overset{\bullet}{\underset{\bullet\bullet\bullet}{\bullet}} \bullet\bullet_\bullet = \bullet \overset{\bullet}{\underset{\bullet\bullet\bullet}{\bullet}} \bullet_\bullet = \bullet$$

- Cuando se multiplica una variable aleatoria por una constante, su esperanza se ve multiplicada por esa constante.

  •[••] = • ••.•. = • • •.•. = ••[•]

  ••• • •[•]• = 0

  Entonces:

  ••• • •[•]• = • (•. • •[•])•. = • •.•. • • •[•]•. = • •.•. • •[•] • •.

  = •[•] • •[•] = 0

- Si *X e Y son* dos variables aleatorias, •[• + •] = •[•] + •[•]

  En general, •[•• + •] = ••[•] + •

Ejemplo 4.4:

En un lote de producción de láminas de Drywall se presenta una probabilidad de obtener lotes defectuosos de la siguiente forma: 0,25 presentan 5 lotes defectuosos, una probabilidad de 0,20 presentan 10, de 0,15 presentan 15, y por ultimo 0,1 presentan 20 lotes defectuosos. Si se considera la variable aleatoria como productos defectuosos. ¿Cuál es su esperanza?

La variable aleatoria X toma los valores de 10, 5, 15, 20 con las probabilidades a que les corresponde 0.2, 0.25, 0.15, 0.1, esto es una variable aleatoria de tipo discreta. Y su esperanza matemática es:

•[•] = 10 (0.25) + 5(0.25) + 20(0.1) + 15(0.15) = 2 + 1.25 + 2 + 2.25 = 7.5 •
8 lotes defectuosos

4.3.2. El Valor Esperado En Variables Aleatorias Continuas

Consideremos una variable aleatoria continua X. La media o esperanza de X se representa mediante la siguiente expresión:

• [•] = • • • • (•) • •

Donde • son los valores que toman la variable y • (•) es la función de densidad.

Propiedades:

- Dada una variable aleatoria continua que toma el valor de C, es decir, la variable es constante. Su esperanza es esa misma constante.

  • [•] = • • • • (•) • • = • • • (•) • • = •

  Teniendo en cuenta que:

  • • (•) • • = 1

- Cuando una variable aleatoria continua se multiplica por una constante, la media también se multiplica por dicha constante.

  • [• •] = • • • • • • (•) • • = • • • • (•) • • = • • • [•]

  • • • • • [•] • = 0

  Entonces:

  • • • • • [•] • = • (• • • [•]) • • (•) • • = • • • (•) • • • • • [•] • • (•) • •

  = • • • • (•) • • • • [•] • • (•) • • = • [•] • • [•] = 0

- Sean • e • dos variables aleatorias continuas, entonces:

  • [• + •] = • [•] + • [•]. En general, • [• • + •] = • • [•] + •

Ejemplo 4.5

Se hace un experimento para determinar el tiempo de duración de una batería de 1.5 voltios, la cual se puede describir por una variable continua *x* mediante la siguiente función:

•(•) = • •3•(•˙ • 1) •ñ•• 0 • • • 1 •
0 •• ••••

Hallar valor esperado de duración de la batería de 1.5 voltios.

Entonces,

•(•) = • • • •(•)•• = • • •(•3•)(•˙ • 1)••

= • (•3•˙ + 3•˙)•• = • •3•˙•• + • 3•˙•• = = •• $\frac{3•˙}{5}$ + •˙• = 0,4

El resultado muestra que en promedio, un repuesto electrónico tiene una vida promedio de 0,4 años, es decir:

0,4 •ñ•• $\frac{•• \; •••••}{• \; •ñ•}$ = 4,8 • •••• • 5 • ••••

Ejemplo 4.6.

El bienestar familiar colombiano, realiza un estudio para evaluar el índice de masa corporal de un grupo de niñas que oscilan en edades entre 13 y 14 años, encontrando que el IMC es una variable aleatoria que describe la siguiente función:

•(•) = •$\frac{1}{0}$ 19 • • • 20 •
•• ••••

El Índice de Masa Corporal (IMC) es una medida de asociación entre el peso y la talla de un individuo.

Hallar la media del índice de masa corporal del grupo de niñas.

•(•) = • • • •(•)•• = • • ••

= •$\frac{•}{2}$• = 19,5

El índice de masa corporal es IMC = $\frac{•••• (••)}{•••••••(•)}$

De acuerdo a la clasificación realizada por la OMS (Organización mundial de la salud), el índice de masa corporal es adecuado. Es decir las niñas tienen un peso apropiado en relación con su estatura.

4.4. VARIANZA EN VARIABLES ALEATORIAS

La varianza de una variable aleatoria es una característica numérica que proporciona una idea de la dispersión de la variable aleatoria respecto de su esperanza. Por lo cual es un parámetro de dispersión.

La varianza es:

•(•) = •[(• • •[•])^2]

En otras palabras, la varianza es el promedio de las desviaciones cuadráticas de los diferentes valores que puede tomar la variable respecto de su valor medio teórico o esperanza.

Una de las características de la varianza es que se expresa en unidades cuadráticas respecto de las unidades originales de la variable. Un parámetro de dispersión derivado de la varianza y que tiene las mismas unidades de la variable aleatoria es la desviación típica, que se define como la raíz cuadrada de la varianza.

$$\bullet_{\bullet} = \bullet\,\overline{\bullet[(\bullet\ \bullet\ \bullet[\bullet])^{2}]}$$

Propiedades de la varianza

1. V(X) • 0
2. • (• · •) = •$^{\bullet}$ · Var(•) para todo número real k.
3. • (•) = 0 para todo número real k.
4. •(• · • + •) = •$^{\bullet}$ · ••• (•) para todo par de números reales • y •.
5. • (• + •) = •(•) + • (•) únicamente en el caso que • e • sean independientes.

4.4.1. VARIANZA EN VARIABLES ALEATORIAS DISCRETAS

La varianza es definida como la esperanza del cuadrado de la diferencia entre la variable y su esperanza:

$$\bullet(\bullet) = \bullet\,[(\bullet - \bullet\,[\bullet])^{\bullet}]$$

Teniendo en cuenta la definición de esperanza, la varianza queda:

$$\bullet(\bullet) = \sum_{\bullet\bullet\bullet}^{\bullet} (\bullet\ \bullet\ \bullet[\bullet])^{\bullet}\bullet_{\bullet}$$

Desarrollando el cuadrado llegamos a la siguiente expresión:

$$\bullet(\bullet) = \sum_{\bullet\bullet\bullet}^{\bullet} \bullet\,(\bullet_{\bullet} \bullet \bullet[\bullet])^{\bullet}\bullet_{\bullet} = \sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}^{\bullet}\bullet_{\bullet} + \sum_{\bullet\bullet\bullet}^{\bullet} \bullet[\bullet]^{\bullet}\bullet_{\bullet} \bullet 2 \sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet\bullet[\bullet]\bullet\bullet_{\bullet}$$

$$= \sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}^{\bullet}\bullet_{\bullet} + \bullet[\bullet]^{\bullet} \sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet} \bullet 2\bullet[\bullet] \sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet_{\bullet} = \bullet[\bullet^{\bullet}] + \bullet[\bullet]^{\bullet} \bullet 2\bullet[\bullet]^{\bullet}$$

$$= \bullet[\bullet^{\bullet}] \bullet \bullet[\bullet]^{\bullet}$$

Finalmente,

$$\bullet(\bullet) = \bullet[\bullet^{\bullet}] \bullet \bullet[\bullet]^{\bullet}$$

$$\bullet^{\bullet} = \sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}^{\bullet}\bullet\bullet_{\bullet} \bullet \bullet^{\bullet} = \sum_{\bullet\bullet\bullet}^{\bullet} (\bullet_{\bullet} \bullet \bullet)^{\bullet}\bullet\bullet_{\bullet}$$

Ejemplo 4.7**:**

Retomando el ejemplo 4.1 se le halla la varianza a la variable aleatoria de la siguiente manera:

Sabiendo que $\bullet(\bullet) = \bullet$ lotes defectuosos;

La variable aleatoria X toma los valores de 10, 5, 15, 20 con las probabilidades a que les corresponde 0.2, 0.25, 0.15, 0.1

$$\bullet^{\bullet} = \sum_{\bullet\bullet\bullet}^{\bullet} ((\bullet\bullet \bullet \bullet)^{\bullet}\bullet\bullet.\bullet + (\bullet \bullet \bullet)^{\bullet}\bullet\bullet.\bullet\bullet + (\bullet\bullet \bullet \bullet)^{\bullet}\bullet\bullet.\bullet\bullet + (\bullet\bullet \bullet \bullet)^{\bullet}\bullet\bullet.\bullet\bullet)$$

$$\bullet^{\bullet} = \bullet\bullet.\bullet$$

4.5. DESVIACIÓN TÍPICA DE VARIABLES ALEATORIAS DISCRETAS

Se llama desviación típica o estándar a la raíz cuadrada positiva de la varianza

$$\bullet = \sqrt{\bullet(\bullet)}$$

Para determinar la varianza hemos elevado los datos al cuadrado, esto implica que las unidades se elevan al cuadrado. La desviación típica consigue tener las desviaciones en la misma unidad que los datos.

• = • • •⁚••. ••˙

Ejemplo 4.8:

Del ejemplo 4.1 podemos hallar la desviación estándar del número de lotes defectuosos de láminas de Drywall que se pueden encontrar.

Reemplazando en la ecuación:

• = •[(•˙••.••) • •˙] + [(10˙••.•) • •˙] + [(••˙••.•) • •˙] + [(10˙••.•) • •˙]

• = •.••

En la distribución de variable discreta se define el momento de orden *k* centrado en el origen como •. = •(•˙), que para variable discreta será:

•. = •(•˙) = • •⁚•(• = •.) = • •⁚•.

Se define el momento de orden • centrado en la media como •. = •[(• • •)˙], que para variable discretas será:

•. = •((• • •)˙) = • (•. • •)˙•(• = •.) = • (•. • •)˙•.

4.5.1. VARIANZA EN VARIABLES ALEATORIAS CONTINUAS

Se define como la esperanza del cuadrado de la diferencia entre la variable y su esperanza: así: $\bullet(\bullet) = \bullet[(\bullet \bullet \bullet[\bullet])^{\bullet}]$, teniendo en cuenta la definición de esperanza:

$$\bullet[\bullet] = \bullet[(\bullet \bullet \bullet[\bullet])^{\bullet}] = \bullet_{\bullet\bullet}^{\bullet} (\bullet \bullet \bullet[\bullet])^{\bullet} \bullet(\bullet)\bullet\bullet$$

Desarrollando el cuadrado queda la siguiente expresión:

$$\bullet[\bullet] = \bullet[\bullet^{\bullet}] \bullet \bullet[\bullet]^{\bullet} = \bullet_{\bullet\bullet}^{\bullet} \bullet^{\bullet}\bullet\bullet(\bullet)\bullet\bullet \bullet \bullet\bullet_{\bullet\bullet}^{\bullet} \bullet\bullet\bullet(\bullet)\bullet\bullet\bullet^{\bullet}$$

Retomando el ejemplo 4.6, hallar la varianza.

El bienestar familiar colombiano, realiza un estudio para evaluar el índice de masa corporal de un grupo de niñas que oscilan en edades entre 13 y 14 años, encontrando que el IMC es una variable aleatoria que describe la siguiente función:

$$\bullet(\bullet) = \bullet\begin{matrix} 1 & 19 \bullet \bullet \bullet 20 \\ 0 & \bullet\bullet \bullet\bullet\bullet\bullet \end{matrix} \bullet$$

Para calcular la varianza es necesario calcular $\bullet[\bullet]$ y $\bullet[\bullet^{\bullet}]$

$$\bullet[\bullet] = \bullet[\bullet^{\bullet}] \bullet \bullet[\bullet]^{\bullet} = \bullet_{\bullet\bullet}^{\bullet\bullet} \bullet^{\bullet}\bullet\bullet(\bullet)\bullet\bullet \bullet \bullet\bullet_{\bullet\bullet}^{\bullet\bullet} \bullet\bullet\bullet(\bullet)\bullet\bullet\bullet^{\bullet}$$

$$= \frac{\bullet^{\bullet}}{3}\bullet\Big|_{\bullet\bullet}^{\bullet\bullet} \bullet \bullet\frac{\bullet^{\bullet}}{2}\bullet\Big|_{\bullet\bullet}^{\bullet\bullet}\bullet^{\bullet} = 380{,}3333 \bullet 380{,}25 = 0{,}0833$$

4.5.2. DESVIACIÓN TÍPICA EN VARIABLES ALEATORIAS CONTINUAS

Se llama desviación típica o estándar a la raíz cuadrada positiva de la varianza

$$\bullet = \bullet\overline{\bullet(\bullet)}$$

Y se define por la siguiente expresión:

$$\sigma = \sqrt{\int_{-\infty}^{\infty} x^2 f(x)\,dx - \left(\int_{-\infty}^{\infty} x f(x)\,dx\right)^2}$$

Del ejemplo anterior, la desviación estándar es igual a:

$$\sigma = \sqrt{0{,}0833} = 0{,}2886$$

4.6. FUNCIÓN DE PROBABILIDAD $f(x)$ DE UNA VARIABLE ALEATORIA

Una función de probabilidad es una función (x) cuyo dominio es un intervalo $(-\infty, \infty)$ y que cumple con las siguientes propiedades:

- $f(x) \geq 0$ para toda x
- $\int_{-\infty}^{\infty} f(x)\,dx = 1$

La función de probabilidad puede reflejar la frecuencia con la que se determinan los valores de cierta población o variable.

4.6.1. Función De Probabilidad $f(x)$ De Una Variable Aleatoria Discreta

Una variable aleatoria discreta X, el cual toma valores $x_1, x_2, \ldots, x_n$. Supongamos que conocemos la probabilidad de que la variable X tome dichos valores, es decir, se conoce que

$P(X=x_1) = p_1$,

$P(X=x_2) = p_2$,

$P(X=x_3) = p_3, \ldots, p(X=x_n) = p_n$

En general $p(X=x_i) = p_i$

La función de probabilidad *f(x)* de la variable aleatoria X es aquella que se asigna a cada valor x_i de la variable su correspondiente probabilidad p_i.

•: • • •   •. •. • 0, •• = 1,2, ... , •
•
• • •(•) = •(• = •)   ••. • •. = 1
•••

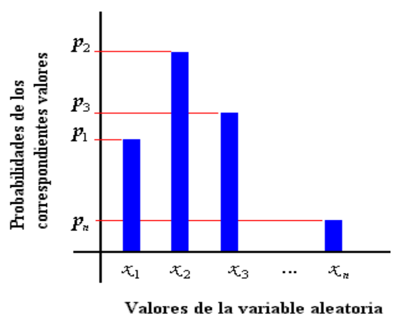
Tabla 4.1 Valores de Función de Probabilidad de una Variable Aleatoria Discreta

X	$\mathbf{x_1}$	$\mathbf{x_2}$	**…**	$\mathbf{x_n}$
$\mathbf{p_i=p(X=x_i)}$	p_1	p_2	…	p_n

Fuente: Los Autores

La representación gráfica más usual de la función de probabilidad es un diagrama de barras no acumulativo.

Figura. 4.1 Función de Probabilidad de una Variable Aleatoria Discreta



Fuente: Los Autores

En otras palabras, se asocia a cada valor de la variable aleatoria discreta la probabilidad acumulada hasta ese valor (la probabilidad de que la variable aleatoria tome valores menores o iguales a x_i).

Propiedades:

F(x) es una probabilidad; 0 • F(x) • 1

F(x) = 0 para todo $\mathbf{x<x_1}$

F(x) = 1 para todo x • x_n

Es constante en cada intervalo [•.,••...)

Es creciente

•(• < • • •) = •(•)••(•)

Podemos expresar la función de distribución de la siguiente forma:

Si una variable aleatoria X tiene como función de probabilidad:

Tabla. 4.2 Valores de una Función de distribución de Probabilidad Acumulada

X	x_1 x_2 … x_n
F(x)= p(X=x)	p_1 p_2 … p_n

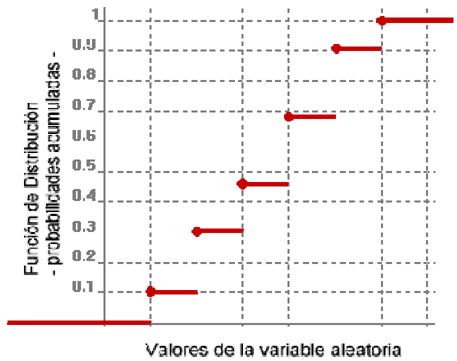
Fuente: Los Autores

Su función de distribución es:

$$F(x) = \begin{cases} 0 & si\ x < x_1 \\ p_1 & si\ x_1 \leq x < x_2 \\ p_1 + p_2 & si\ x_2 \leq x < x_3 \\ \cdots & \cdots \\ p_1 + p_2 + \ldots + p_{n-1} & si\ x_{n-1} \leq x < x_n \\ 1 & si\ x \geq x_n \end{cases}$$

Su representación gráfica tiene forma escalonada, donde los saltos coinciden con la probabilidad •ᵢ correspondiente a los valores •ᵢ de la variable •.

Figura. 4.2 Función de distribución de Probabilidad Acumulada



Fuente: Los Autores

Ejemplo 4.9.

El departamento de mercadeo de una empresa de mensajería, realiza la trazabilidad de los paquetes enviados que no llegan a tiempo, lo cual ocasiona quejas por parte de los clientes, quienes confían en la llegada oportuna de los paquetes enviados. De acuerdo a la información recopilada se encuentra que:
En el mes de abril de este año, un total de 20 paquetes no llegaron en el tiempo convenido con el cliente. El cuadro siguiente detalla el número de paquetes que no fueron entregados a tiempo en cada semana.

Tabla 4.3 Numero de Paquetes que no fueron entregados a tiempo en semanas.

Semana	1	2	3	4
No. De Paquetes que no fueron entregados a tiempos	3	9	4	4

Fuente: Los Autores

La meta del departamento de mercadeo es reducir el 5% de las quejas presentadas por llegadas el incumplimiento a los clientes en cuanto a los tiempos de entrega de los paquetes enviados.

El personal de mercadeo está haciendo todas las gestiones posibles para que el próximo mes, no se presente la misma situación del año pasado, en el cual los meses de abril y mayo presentaron igual número de quejas por incumplimiento en los tiempos de entrega de los paquetes enviados por los clientes.

El jefe del departamento quiere conocer la probabilidad que se presenten quejas por parte de los clientes en cada una de las semanas del mes de mayo. Suponiendo que la variable se comporte igual que el mes de abril.

Para estimar cada una de las probabilidades se aplica la siguiente fórmula:

$$\bullet(\bullet = \bullet\bullet\bullet\bullet\bullet\bullet\bullet) = \frac{\bullet\bullet\ \bullet\bullet \qquad \bullet\bullet\bullet\ \bullet\bullet \qquad \bullet\bullet\bullet\bullet\bullet\ \bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet}{\bullet\bullet\bullet\bullet\bullet\ \bullet\bullet \qquad \bullet\bullet\bullet\ \bullet\bullet \qquad \bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet}$$

$$\bullet(\bullet = \bullet\bullet\bullet\bullet\bullet\bullet\ 1) = \frac{3}{20}$$

$$\bullet(\bullet = \bullet\bullet\bullet\bullet\bullet\bullet\ 2) = \frac{\bullet}{\bullet\bullet}$$

$$\bullet(\bullet = \bullet\bullet\bullet\bullet\bullet\bullet\ 3) = \frac{4}{20}$$

$$\bullet(\bullet = \bullet\bullet\bullet\bullet\bullet\bullet\ 4) = \frac{4}{20}$$

En resumen:

Tabla 4.4 Probabilidad de quejas en semanas.

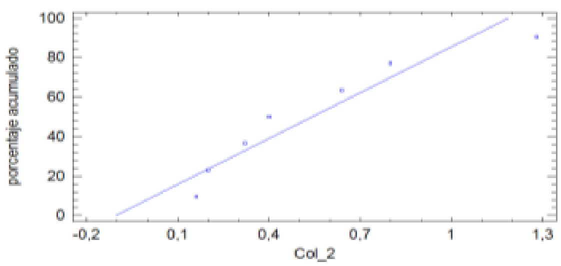
Semana	1	2	3	4
Probabilidad que se presenten quejas	0.15	0.45	0.2	0.2

Fuente: Los Autores

4.6.2. FUNCIÓN DE PROBABILIDAD DE UNA VARIABLE ALEATORIA CONTINUA

La función de probabilidad •(•) nos brinda la información que es llamada densidad de probabilidad, lo que quiere decir que el área que se encuentra por debajo de la función entre dos valores de la variable [•, •] es la misma probabilidad de que ésta pueda tomar cualquier valor entre los mismos, como se muestra en la siguiente figura.

Figura 4.3 Función de Probabilidad de una Variable Aleatoria Continua



Fuente: Los Autores

La probabilidad de que la variable • pueda tomar cualquier valor entre • y • se calcula de la siguiente forma:

•(• • • • •) = • •(•)••

Si el área que se encuentra debajo de la función sirve para la determinación de la probabilidad en un intervalo de valores, por lo tanto se debe cumplir que el área total por debajo de ella debe ser 1;

• •(•)•• = 1

Y en consecuencia de que • sea continúa, o sea que cualquier valor • de la variable su probabilidad es nula; •(• = •) = 0. Este resultado se deriva de:

$$•(• = •) = • \; •(•)•• = 0$$

En resumen, una función •(•) es una función de probabilidad de una variable aleatoria continua si cumple con las siguientes condiciones:

- Es positiva en todo su dominio: 0 • •(•) • 1
- La probabilidad •(• • •(•) • •) corresponde al área bajo la gráfica, comprendida en los valores *a* y *b* de la variable *x*.

  $•(• • • • •) = • \; •(•)••$
- El área total de la gráfica de • y el eje X, equivale a 1. Es decir,

  $• \; •(•)•• = 1$
- Permite obtener F(x) como área bajo la gráfica hasta el valor de *x.*

Ejemplo 4.10:

Hallar la probabilidad que el precio de una libra de café varíe entre 1,0 y 2,0 dólares, de acuerdo a la siguiente función de probabilidad

• (•)=Probabilidad de una libra de cafe.

$$•(•) = • \begin{matrix} 2• \; • \; 2 & 1{,}0 \; • \; • \; • \; 2{,}0 \\ 0 & •• \; •••• \end{matrix} \; •$$

$$• = • \; (2• \; • \; 2)•• = |• \; • \; 2•| = 1$$

EJERCICIOS PROPUESTOS
CAPÍTULO N° 4.

Preguntas Propuestas.

Responda las siguiente preguntas, marque con una X si está es Verdadero (V) ó Falso (F)

1. Se define que una variable aleatoria es la que se encuentra asociada a un experimento aleatorio, y a este se le asocia un valor numérico a cada valor del experimento, estos valores pueden ser de dos tipos tanto positivos como negativos, existen dos tipos de variable aleatoria; los cuales son discreta y continua. ________

2. Se denomina *variable discreta* si únicamente toma valores finitos o numerables; y *variable continua* cuando toma valores infinitos ó no numerables. __________

3. Se denomina función de probabilidad de una variable aleatoria cuando a este se le es asignada una probabilidad. Cuando se especifican los diferentes valores de la variable aleatoria y sus diferentes probabilidades, de esta forma se construye un modelo de distribución de probabilidad. ___________

Ejercicios propuestos.

1. Dada la ley de probabilidad discreta P(X=i) = i/120, *i*=1,2,3,…,10 calcular: P $(2 \leq X \leq 7)$= 0 y P $(X^2 \leq 50)$. Calcular también la esperanza y la varianza de la variable X

2. La siguiente tabla contiene la distribución de probabilidad del número de pacientes diarios que son atendidos en un hospital psiquiátrico.

Tabla 4.5 Probabilidad de Pacientes Atendidos en el Hospital Psiquiátrico

PACIENTES ATENDIDOS EN EL HOSPITAL PSIQUIATRICO	
Pacientes atendidos diariamente (y)	Probabilidad P(y)
0	0,05
1	0,10
2	0,25
3	0,15
4	0,30
5	0,15

Fuente: Los Autores

Calcula el número de pacientes atendidos por día.

3. En un estudio realizado para determinar la relación del uso del internet con el riesgo de padecer depresión entre jóvenes de 18 a 25 años, se determinó que esta relación se describe a través de una función de probabilidad:

$$\begin{cases} \frac{2}{5}(x+2) & 0 \leq x \leq 1 \\ 0 & \text{en cualquier punto} \end{cases}$$

Determinar el número esperado de jóvenes que sufren depresión.

4. La velocidad del arrastre de un colido genera en una fuente fluvial contaminación debido a la acumulación de sedimentos en el fondo del mismo, la probabilidad de los sedimentos en su totalidad está dada por la siguiente función:

$$\begin{cases} (4x+1) & 0 \leq x \leq 1 \\ 0 & \text{en cualquier punto} \end{cases}$$

Hallar la probabilidad de que se dispersen totalmente los sedimentos.

5. La tasa de mortalidad de una ciudad se encuentra dada por la siguiente distribución de probabilidad.

$$\bullet \begin{array}{ll} (e^{\bullet} \bullet e + 2) & 0 \bullet x \bullet 1 \\ 0 & \text{en cualquier punto} \end{array} \bullet$$

   a. Calcula el valor esperado y varianza del tiempo anual
   b. Si cada familia de la persona fallecida se le cancelan \$2000 dólares por seguro de vida anual. Calcular el valor esperado y varianza.
   c. Cree que el costo excedería con frecuencia los \$2.000.000

6. Los costos de combustible para una empresa se manejan con la siguiente distribución,

$$\bullet \begin{array}{ll} (6x^{\bullet} \bullet 1) & 0 \bullet x \bullet 1 \\ 0 & \text{en cualquier punto} \end{array} \bullet$$

   Si en la empresa existen nueve vehículos que consumen 55gl - diarios por vehículo, calcule el valor esperado y la varianza del costo de galones mensuales.

7. La producción de una empresa de concreto se encuentra dada por la siguiente distribución en (metros cúbicos – minutos):

$$\bullet \begin{array}{ll} (\bullet 2xe^{\bullet\bullet} + e) & 0 \bullet x \bullet 1 \\ 0 & \text{en cualquier punto} \end{array} \bullet$$

   El costo de un metro cubico es de:
   Agua \$11m3
   Grava \$250 /kg
   Arena \$ 125 /kg
   Cemento \$ 450/kg

¿Cuál sería el punto de equilibrio para la empresa, si además se deben de saber cuántos metros se deben de producir sabiendo que la hora hombre es equivalente a $1300 y se necesitan 5 operarios para el proceso, y los imprevistos son del 10% de lo producido?

ESTUDIO DE CASOS

CASO 1

En una empresa metalmecánica fabrican piezas en acero inoxidable, estas piezas de acero inoxidable deben cumplir las especificaciones técnicas del cliente. por lo cual se debe asegurar las medidas críticas que se encuentren dentro de los rangos de tolerancia.

Para fabricar la pieza, se realizan dos procesos básicos, torneado y fresado. El fresado es el proceso en el cual ocurren el 80% de los reprocesos. Por lo cual la empresa realizó un estudio para establecer las causas de las mismas.

A partir del análisis de causas se llegó a la conclusión que los reprocesos son causados por fallas en las máquinas. A continuación se muestra la cantidad de piezas que se fabrican en el mes:

Descripción	Enero	Febrero	Marzo	Abril	Mayo	Junio
Total Piezas Fabricadas	94	93	136	140	160	148

Otra conclusión derivada del estudio, es que las no conformidades se generan cuando la operación se realiza a altas velocidades por lo que cuando se realizan las piezas a estas velocidades la probabilidad que una pieza no cumpla las especificaciones es igual a 0.01 del 70% de las piezas que se fabriquen. ,

Para los meses de Julio, agosto, septiembre y octubre, la empresa debe cumplir con las demandas de los clientes los cuales son:

Mes	Demanda
Julio	215
Agosto	180
Septiembre	145
Octubre	123

Para cumplir con la demanda de los clientes, las maquinas deben operar a altas velocidades. Teniendo en cuenta lo anterior:

1. Cuál es la probabilidad que se incrementen los reprocesos en un 10%
2. Genere alternativas para disminuir los reprocesos

CASO 2

Un partido político realizo un estudio con el fin de evaluar las preferencias de las personas con el potencial de votar en las épocas de elecciones.

Dicho estudio es realizado a 200 personas en 10 de las principales ciudades de Colombia. En el estudio se evalúan aspectos decisivos por los electores, tales como: Educación, Salud, Seguridad, Capacidad de generar empleo.

Los resultados fueron los siguientes:
Electores que se inclinan por un gobierno donde la prioridad es la capacidad de generar empleo: 60%
Electores que se inclinan por un gobierno donde la prioridad es la educación: 60%
Electores que se inclinan por un gobierno donde la prioridad es la salud: 60%
Electores que se inclinan por un gobierno donde la prioridad es la seguridad: 60%

Además el 30% de las personas con potencial de votar no están seguras de sufragar.
El 40% no sufragan porque consideran no votar una alternativa para manifestar que no están de acuerdo son ningún partido político y el 25% ejercen su derecho al voto.

Otro aspecto muy relevante es que el 30% que no están seguras por quién votar, el 25% son jóvenes.
Para alcanzar el porcentaje de las votaciones necesaria para ganar las elecciones, el partido político decide fortalecer su proyecto de generación de empleo y llegar a la población de personas que ejercen su derecho al voto y los jóvenes que no están seguros por quién votar. Teniendo en cuenta lo anterior:

1. Cree usted que la estrategia a implementar por el partido político es la mejor. Argumente su respuesta

2. Proponga una estrategia que asegure mayor probabilidad de ganar las elecciones

CAPÍTULO 5

DISTRIBUCIONES DE PROBABILIDAD

INTRODUCCION

La distribución de probabilidad es una función que asigna a cada suceso de una variable aleatoria la probabilidad de ocurrencia del suceso.

Las distribuciones de probabilidad describe como se distribuye las probabilidades de los diferentes valores de una variable aleatoria, la aleatoriedad se asocia con la ocurrencia de ciertos eventos cuando se lleva a cabo un experimento. Estas se pueden representar gráficamente:

1. Cuando la variable aleatoria es discreta, la gráfica de la función de probabilidad puede construirse usando segmentos de retas verticales. Los valores de la variable aleatoria se localizan en el eje horizontal y las probabilidades en el eje vertical, en cada valor se construye un segmento de recta vertical de altura igual a la probabilidad de la variable aleatoria. La suma de las longitudes verticales debe ser igual a uno.

2. A través de un histograma de probabilidad en donde los rectángulos están dibujados de tal forma que sus bases, con el mismo ancho, están centradas en cada valor de x de X, y sus alturas son iguales a la correspondientes probabilidades dadas por $f(x)$, puesto que cada base tiene un ancho igual a uno.

A través de la distribución de probabilidad se puede describir como se espera que se comporte los resultados de una variable. Por lo cual pueden representar modelos que sirven para realizar inferencias y tomar decisiones en condiciones de incertidumbre.

Muchos conjuntos de datos se pueden aproximar, con gran exactitud, utilizando determinadas distribuciones probabilísticas; los resultados de éstas se pueden utilizar para analizar datos estadísticos.

En este capítulo se trataran las diferentes distribuciones de probabilidad para las variables aleatorias discretas y continuas.

COMPETENCIAS:

- Interpretación y aplicación de las distribuciones de probabilidad de variables aleatorias continuas y discretas
- Interpretacion e identificación de tablas de probabilidad de las diferentes distribuciones de probabilidad para determinar la probabilidad de una variable aleatoria
- Interpretación y Calculo de los teoremas para determinar el valor esperado de una variable aleatoria con una distribución de probabilidad.
- Utilizacion conceptos básicos de distribuciones de probabilidad
- Diseño de experimentos aleatorios (de las ciencias físicas, naturales o sociales) para estudiar un problema o pregunta.
- Proposicion de inferencias a partir del estudio de las distribuciones de probabilidad.

5. DISTRIBUCION DE PROBABILIDAD

5.1 DISTRIBUCIÓN DE PROBABILIDAD PARA UNA VARIABLE DISCRETA

Este tipo de distribuciones de probabilidad es un listado mutuamente excluyente de todos los posibles resultados, para este tipo de variables, tanto que se dice que la probabilidad particular se encuentra asociada con cada uno de los resultados-

Ejemplo:

Mauricio ha lanzado dado, cuál es la distribución de probabilidad para hallar el resultado que obtuvo Mauricio.

Distribución de probabilidad de un dado	**1**	**2**	**3**	**4**	**5**	**6**
Probabilidad	1/6	1/6	1/6	1/6	1/6	1/6

Entonces:

La probabilidad que se obtenga 3 como resultado es 1/6

Usando la regla de la adición para eventos mutuamente excluyentes, la probabilidad de una cara impar es:

P (impar) = P(1)+P(3)+P(4)= 1/6+1/6+1/6 = 2/6

La probabilidad de que la cara sea mayor que 6 es P(• > 6)=0, X se define como el posible reslutado.

En la distribución de probabilidad discreta se cumple que:

- 0 • •(•) • 1 para toda *y*
- •. •(•) = 1, donde la sumatoria es sobre todos los valores posibles de la variable *y*, con una probabilidad diferente de cero.

5.1.1 MODELOS DE PROBABILIDAD DE VARIABLES ALEATORIAS DISCRETAS

Dentro de los modelos de probabilidad, correspondiente a variables aleatorias discretas, las de mayor aplicación se tienen: Bernoulli, Binomial, Poisson, Exponencial, Multinomial e hipergeométrico. Mientras que los modelos de probabilidad correspondiente a variables aleatoria continuos se considera el modelo normal estandarizado.

5.1.1.1 Distribución Bernoulli

Un experimento de Bernoulli tiene uno de dos resultados mutuamente excluyentes, que generalmente se denotan S (éxito) y F (fracaso). Por ejemplo, al seleccionar un producto con el fin de verificar que cumpla con las especificaciones de calidad, se lleva a control de calidad, después de efectuar las pruebas necesarias puede ocurrir que este cumpla o no cumpla las especificaciones. El espacio muestral E, para este experimento Bernoulli consta de dos resultados E= {cumple, no cumple}.

Supongamos que ***p*** es la probabilidad de observar un éxito en un ensayo Bernoulli, es decir que el producto cumpla con las especificaciones, donde 0 • -• • 1.

La distribución de probabilidad de la variable aleatoria en Bernoulli, está dada por:

•(• = 1) = •, •(• = 0) = 1 – • = •

Donde x el resultado que se obtiene a partir de las pruebas de calidad, los cuales pueden ser que el producto cumpla o que no cumpla las especificaciones.

La distribución de probabilidad de la variable aleatoria de Bernoulli *X* está dada por:

X ~ Bernoulli (*p*), *P(X=x)* = •˙•˙˙˙, • = 0,1

Si • variable de Bernoulli (*p*), entonces

E(X)=*p*

E(X) = 0·*q* + 1· *p* = *p*

•˙ (X) = E[(X • E(X))˙] = (0 • •)˙• + (1 • •)˙• =

= •˙• + •˙• = ••(• + •) = ••

•˙ (X) = ••

Propiedades

1. Función de distribución

	0	•• • < 0
•(•) = •(• • •) = •	1 • •	•• 0 • • < 1•
	1	•• • • 1

Ejemplo 5.1:

Si X variable de Bernoulli (1/3) • P(*x*) = $(1/3)^{x}(2/3)^{1-x}$

Entonces *p* = 1/3 y *q* = 2/3, luego

E(X) = 1/3 y •˙(X) = 2/9

Una variable aleatoria de Bernoulli, por sí sola, tiene poco interés en las aplicaciones de ingeniería y ciencias. Pero cuando estas son aplicables con una serie de pruebas de Bernoulli, produce varias distribuciones de probabilidad discretas bien conocidas y útiles.

5.1.1.2 Distribución Uniforme

La distribución uniforme es la que corresponde a una variable que toma todos sus valores, x_1, x_2..., x_n, con una probabilidad igual; el espacio muestral debe ser finito. Si la variable tiene k posibles valores, la función de probabilidad sería:

$$\bullet\bullet\ \ \bullet\ \bullet\ \ \bullet(\bullet,\bullet) = \frac{1}{\bullet}$$

En donde k es un parámetro de la distribución (un parámetro es un valor que sirve para determinar la función de probabilidad o densidad de una variable aleatoria)
La media y la varianza de la variable uniforme se calculan por las expresiones:

$$\bullet(\bullet) = \sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet\bullet(\bullet = \bullet_{\bullet})$$

$$\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet\frac{1}{\bullet} = \frac{1}{\bullet}\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}$$

$$\bullet^{\bullet}(\bullet) = \frac{1}{\bullet}\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}^{\bullet}\ \bullet\ \bullet\frac{1}{\bullet}\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet^{\bullet}$$

$$\bullet^{\bullet}(\bullet) = \bullet(\bullet^{\bullet})\ \bullet\ \bullet\bullet(\bullet)\bullet^{\bullet} = \sum_{\bullet\bullet\bullet}^{\bullet} \bullet^{\bullet}\bullet\bullet(\bullet = \bullet_{\bullet})\ \bullet\ \bullet\frac{1}{\bullet}\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet^{\bullet}$$

$$= \sum_{\bullet\bullet\bullet}^{\bullet} \bullet^{\bullet}\bullet\frac{1}{\bullet}\ \bullet\ \bullet\frac{1}{\bullet}\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet^{\bullet} = \frac{1}{\bullet}\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}^{\bullet}\ \bullet\ \bullet\frac{1}{\bullet}\sum_{\bullet\bullet\bullet}^{\bullet} \bullet_{\bullet}\bullet^{\bullet}$$

Propiedades

1. Función de Distribución**:**

$$\bullet(\bullet) = \bullet(\bullet\ \bullet\ \bullet) = \sum_{\bullet.\bullet\bullet} \bullet(\bullet = \bullet.)$$

5.1.1.3 Distribución Binomial

Esta distribución corresponde a una distribución de variable aleatoria discreta. Para calcular esta distribución se hace necesario recordar el binomio de Newton y el triángulo de Pascal.
Binomio de Newton: Puede describirse de la siguiente manera:

$$(\bullet + \bullet)^{\bullet} = \bullet^{\bullet} + \bullet\bullet^{\bullet\bullet\bullet}\bullet + \frac{\bullet(\bullet\ \bullet\ 1)}{1 \times 2}\bullet^{\bullet\bullet\bullet}\bullet^{\bullet} + \bullet\ + \bullet^{\bullet}$$

$$(\bullet + \bullet)^{\bullet} = \bullet + \bullet$$

$$(\bullet + \bullet)^{\bullet} = \bullet^{\bullet} + 2\bullet\bullet + \bullet^{\bullet}$$

$$(\bullet + \bullet)^{\bullet} = \bullet^{\bullet} + 3\bullet^{\bullet}\bullet + 3\bullet\bullet^{\bullet} + \bullet^{\bullet}$$

$$(\bullet + \bullet)^{\bullet} = \bullet^{\bullet} + 4\bullet^{\bullet}\bullet + 6\bullet^{\bullet}\bullet^{\bullet} + 4\bullet\bullet^{\bullet} + \bullet^{\bullet}$$

$$(\bullet + \bullet)^{\bullet} = \bullet^{\bullet} + 5\bullet^{\bullet}\bullet + 10\bullet^{\bullet}\bullet^{\bullet} + 10\bullet^{\bullet}\bullet^{\bullet} + 5\bullet\bullet^{\bullet} + \bullet^{\bullet}$$

$$(\bullet + \bullet)^{\bullet} = \bullet^{\bullet} + 6\bullet^{\bullet}\bullet + 15\bullet^{\bullet}\bullet^{\bullet} + 30\bullet^{\bullet}\bullet^{\bullet} + 15\bullet^{\bullet}\bullet^{\bullet} + 6\bullet\bullet^{\bullet} + \bullet^{\bullet}$$

La formula de $(\bullet + \bullet)^{\bullet} = \bullet\binom{\bullet}{\bullet}\bullet\bullet^{\bullet\bullet\bullet}\bullet^{\bullet}$ permite establecer una correspondencia entre los coeficientes binomiales y el numero de combinaciones.

Sea A un suceso que tiene una probabilidad “*p*” de ocurrir. Para observar si en cada experiencia ocurre el suceso A. Supongamos, además, que las repeticiones son independientes, es decir, que después de cada repetición la probabilidad “p” de que ocurra el suceso A no se ha modificado.

Ahora bien, para cada "número de veces que ocurre el suceso A, de probabilidad p, cuando efectuamos n repeticiones independientes de la experiencia", designamos por X a la variable aleatoria entonces X es una variable binomial de parámetros n y p y se representa:

$$X \equiv B(n,p)$$

En el caso de una variable aleatoria binomial la *Función de Probabilidad* viene definida por la ecuación:

$$P_X(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \quad \text{con } x=0,1,2,...,n$$

La *media* y la *varianza* de una variable binomial se determinan mediante:

• (•) = • •

Sabemos que una variable del tipo binomial • (• , •) puede expresarse como una suma de • variables de tipo Bernouilli • (•), por tanto:

• (•) = • (•. + •. + • + •.) = • (•.) + • (•.) + • + • (•.) =

= • + • + • + • = • •

$\sigma^2(X) = n \cdot p \cdot (1-p)$; Donde • = 1 • •

•˙(•) = •˙(•. + •. + • + •.) = •˙(•.) + •˙(•.) + • + •˙(•.) =

= • • + • • + • + • • = • • •

Otra cuestión de interés a estudiar es el caso de la *adición* de distribuciones: la suma de dos variables binomiales independientes con idéntico parámetro p es otra Binomial:

$$B(n_1,p)+B(n_2,p)=B(n_1+n_2,p)$$

Ejemplo 5.2:

Una agencia de viajes vende paquetes de viajes que incluyen: tiquete de aeropuerto-hotel-aeropuerto y viceversa, hospedaje por tres días y dos noches en una habitación doble, y desayuno, de este tipo de paquete vende a 7 personas, todos el mismo día. La probabilidad de que una persona viaje en quince días es de 0.7, determinar la probabilidad que dentro de quince días viajen:

a) Las 7 personas
b) Al menos 4
c) Solo 3
d) Al menos 1

Es una situación en donde las personas tienen dos opciones de ir de viaje en quince días o de no ir de viaje en quince días, esta tiene una probabilidad de éxito (asistir al viaje en quince días) o de fracaso (no asistir al viaje en quince días)

a) Que asistan las 7 personas

P(X=7) è

$$P(X=7)=\binom{7}{7}\cdot 0.7^{7}\cdot(0.3)^{0}=1\cdot 0{,}0823\cdot 1=0{,}0823$$

b) Al menos 4

P(• • 4) è

P(• • 4) = 1 • •(• = 4) = 1 • •(• = 0) • •(• = 1) • •(• = 2) • •(• = 3) =

Se hallan las probabilidades para cada una de las *X*; y son las siguientes

P(• = 0) = 0.00021

P(• = 1) = 0.00357

P(• = 2) = 0.02500

P(• = 3) = 0.09724

P(• • 4) = 1 • •(• = 4) = 1 • 0.00021 • 0.00357 • 0.02500 • 0.09724 = 0.8739

c) Solo 3

P(X=3)

$$P(X=3)=\binom{7}{3}\cdot 0.7^{3}\cdot(0.3)^{7-3}=35\cdot 0.343\cdot 0.0081=0{,}09724$$

d) Al menos 1

P(X=1) è •(• = 1) = 1 • •(• = 0) = 1 • 0.00021 = 0.9997

5.1.1.4 Distribución Multinominal

La distribución multinomial es muy similar a la binomial con la diferencia que cada prueba tiene más de dos posibles resultados mutuamente excluyentes.

Si tenemos que K es un resultados posibles (E_i,i = 1, ... , K) con probabilidades fijas (p_i, i = 1, ... , K), la variable que expresa el número de resultados de cada tipo obtenidos en *n* pruebas realizadas de manera independientes tiene distribución multinomial.

Tabla 5.1 Valores de una Distribución Multinominal

Resultados	No. De ocurrencias en *n* pruebas	Probabilidad
E_1	X_1	P_1
E_2	X_2	P_2
…	…	…
E_k	X_k	P_k
Total	• • •. = • •••	• • •. = 1 •••

Fuente: Los Autores

Una variable aleatoria k-dimensional (••, … , ••) sigue una distribución del parámetro •, ••, … , ••, su función de probabilidad es la siguiente:

•(••, … , ••) = •••…., … , •…. • = •! / (••! …… ••!) •• …… ••

Para •• = 0,1, … , •, con •… •• = •. Se dice que la variable (••, … , ••) es de tipo • (•, ••, … , ••),

La *media* y la *varianza* de una variable k-dimensional se determinan mediante:

•(••) = •••

•(••) = •• (••, … , ••) / •• … = •• / •• • • ••••• • • …

= ••••••(••••• + • + ••••• + • + ••••)•••|……

= •••(•• + • ••)••• = •••

••(•) = •••(1 • ••)

•(••) = •• (••, … , ••) / •• … = •••• • ••• + •••

Se tiene que:

••(•) = ••••• • ••••• •• = •••• • ••• + ••• • ••••

••(•) = •••(1 • ••)

5.1.1.5 Distribución de Poisson

Consideremos X una variable Binomial de parámetros n y p. Cuando n tiende hacia infinito y p tiende hacia 0, manteniéndose constante el valor medio λ= • · •, por lo cual X tiene una distribución de Poisson y la representaremos por:

$$\bullet \equiv \bullet\bullet(\lambda)$$

Para la distribución de Poisson la *Función de Probabilidad* es:

$$P_X(x)=\frac{e^{-\lambda}\cdot\lambda^x}{x!} \quad \text{con } x=0,\ 1,\ 2,\ ...$$

Y la *media* y la *varianza* son:

$$\bullet(\bullet) = \lambda$$

$$\bullet(\bullet) = \overset{\bullet}{\underset{\bullet\bullet\bullet}{\bullet}} \bullet\bullet(\bullet = \bullet) = \overset{\bullet}{\underset{\bullet\bullet\bullet}{\bullet}} \bullet\frac{\lambda^{\bullet}}{\bullet!}\bullet^{\bullet\lambda}$$

$$= \bullet^{\bullet\lambda}\bullet 0 + 1\frac{\lambda}{1!} + 2\frac{\lambda^{\bullet}}{2!} + 3\frac{\lambda^{\bullet}}{3!} + \bullet\bullet =$$

$$= \bullet^{\lambda}\bullet\lambda + \frac{\lambda^{\bullet}}{1!} + \frac{\lambda^{\bullet}}{2!} + \bullet\bullet =$$

$$= \bullet^{\bullet\lambda}\lambda\bullet 1 + \frac{\lambda}{1!} + \frac{\lambda}{2!} + \bullet\bullet =$$

$$= \bullet^{\bullet\lambda}\bullet\lambda\bullet\bullet^{\lambda} = \lambda$$

$$\bullet^{\bullet}(X)=\lambda$$

$$\bullet(\bullet^{\bullet}) = \overset{\bullet}{\underset{\bullet\bullet\bullet}{\bullet}} \bullet^{\bullet}\bullet(\bullet = \bullet) = \overset{\bullet}{\underset{\bullet\bullet\bullet}{\bullet}} \bullet^{\bullet}\frac{\lambda^{\bullet}}{\bullet!}\bullet^{\bullet\lambda} =$$

$$= \bullet^{\bullet\lambda}\bullet 0 + 1\frac{\lambda}{1!} + 2^{\bullet}\frac{\lambda^{\bullet}}{2!} + 3^{\bullet}\frac{\lambda^{\bullet}}{3!} + \bullet\bullet$$

$$= \bullet^{\bullet\lambda}\bullet\lambda + 2\lambda^{\bullet} + \frac{3\lambda^{\bullet}}{2!} + \frac{4\lambda^{\bullet}}{3!} + \bullet\bullet$$

$$= \lambda\bullet^{\bullet\lambda}\bullet 1 + 2\lambda + \frac{3\lambda^{\bullet}}{2!} + \frac{4\lambda^{\bullet}}{3!} + \bullet\bullet$$

$$= \lambda\bullet^{\bullet\lambda}\bullet 1 + (1+1)\lambda + (1+2)\frac{\lambda^{\bullet}}{2!} + (1+3)\frac{\lambda^{\bullet}}{3!} + \bullet\bullet$$

$$= \lambda\bullet^{\bullet\lambda}\bullet\bullet^{\lambda} + \lambda\bullet 1 + \frac{\lambda}{1!} + \frac{\lambda^{\bullet}}{2!} + \bullet\bullet\bullet$$

$$= \lambda\bullet^{\bullet\lambda}\bullet\bullet^{\lambda} + \lambda\bullet^{\lambda}\bullet = \lambda(1+\lambda) = \lambda + \lambda^{\bullet}$$

La *adición* para el caso de la distribución de Poisson se define como: la suma de dos variables de Poisson independientes resultando otra variable de Poisson cuyo parámetro es la suma de parámetros de las variables sumadas:

$$Ps(\lambda_1)+Ps(\lambda_2)=Ps(\lambda_1+\lambda_2)$$

Ejemplo 5.3:

En la fila de un banco el número promedio que llegan las personas a la cola es de 150 por hora, si el banco puede atender máximo un total de 8 personas por minuto, hallar la probabilidad que en un minuto cualquiera llegue mas personas a la cola esperando ser atendida?

El número de personas aleatorio el cual llega a la cola esperando ser atendido en minutos se encuentra dado de Pla siguiente forma, siendo *X* la variable de Poisson de parámetro *m=150/60=2,5,* este sería el número medio de llegadas a la cola del banco en minutos

La probabilidad que se debe hallar es $P(\bullet > 8)$, ya que para que lleguen mas personas a la cola la probabilidad se debe encontrar por encima de la media, entonces,
$\bullet(\bullet > 8) = 1 \bullet \bullet(\bullet \bullet 8)$, siendo $\bullet(\bullet \bullet 8)$ la función de distribución de la variable de parámetro 2.5

$$\bullet(\bullet \bullet 8) = \bullet(\bullet = 0) + \bullet(\bullet = 1) + \bullet(\bullet = 2) + \bullet(\bullet = 3) + \bullet(\bullet = 4) + \bullet(\bullet = 5) + \bullet(\bullet = 6) + \bullet(\bullet = 7)$$

Entonces, aplicando la formula de probabilidad para una distribución de poisson

$$\bullet(\bullet =)\frac{\bullet^{\bullet}}{\bullet!}\bullet^{\bullet\bullet}$$

Obtenemos que:

•(• = 0) = 0,08

•(• = 1) = 0,21

•(• = 2) = 0,26

•(• = 3) = 0,21

•(• = 4) = 0,13

•(• = 5) = 0,07

•(• = 6) = 0,03

•(• = 7) = 0,01

Entonces •(• • •) = • • • • = •

La probabilidad de que sea mayor es **0**

5.1.1.6 Distribución Hipergeométrica

Supongamos una población formada por N individuos (finita) de la que extraemos al azar n de ellos. Comprobamos cuántas veces ocurre un determinado suceso A que tiene una probabilidad p de ocurrir. Si llamamos X al número de individuos de entre los n que hemos tomado al azar que tienen esa característica A entonces • es una variable Hipergeométrica y lo denotaremos por • = •(•, •, •).

La distribución Hipergeométrica es el mismo que el de la Binomial. La única diferencia es que en el caso de la Binomial las repeticiones debían ser independientes y en este caso deben ser dependientes.

X=H(M,n,p)

La *Función de Probabilidad* de una distribución Hipergeométrica es:

$$P_X(x) = \frac{\binom{M \cdot p}{x} \cdot \binom{M \cdot (1-p)}{n-x}}{\binom{M}{n}} \quad \bullet\bullet\bullet\ \bullet = 0,1,2,\ldots,\bullet$$

Y la *media* y la *varianza* son:

$$\bullet(\bullet) = \bullet\bullet$$

$$\bullet(\bullet) = \sum_{\bullet} \bullet\bullet\bullet(\bullet = \bullet) = \sum_{\bullet} \bullet\bullet\frac{\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet}{\bullet\bullet\bullet\bullet} =$$

$$= \frac{1}{\bullet\bullet\bullet\bullet} \sum_{\bullet} \bullet\bullet\frac{\bullet\bullet\bullet(\bullet\bullet\bullet 1)\ldots\ldots(\bullet\bullet\bullet\bullet + 1)\bullet\bullet\bullet\bullet\bullet\bullet\bullet}{\bullet\bullet(\bullet\bullet 1)\ldots\ldots 1} =$$

$$= \frac{\bullet\bullet}{\bullet\bullet\bullet\bullet} \sum_{\bullet} \bullet\frac{\bullet\bullet\bullet 1}{\bullet\bullet 1}\bullet\bullet\frac{\bullet\bullet}{\bullet\bullet\bullet}\bullet = \frac{\bullet\bullet}{\bullet\bullet\bullet\bullet}\bullet\frac{\bullet\bullet 1}{\bullet\bullet 1}\bullet$$

$$= \frac{\bullet\bullet}{\bullet}\bullet\bullet = \bullet\bullet$$

$$\bullet^{\bullet}(\bullet) = \bullet\bullet\bullet\frac{\bullet\bullet\bullet}{\bullet\bullet 1}$$

$$\bullet(\bullet^{\bullet}) = \sum_{\bullet} \bullet^{\bullet}\bullet\bullet(\bullet = \bullet) = \sum_{\bullet} \bullet^{\bullet}\bullet\frac{\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet}{\bullet\bullet\bullet\bullet}$$

$$= \sum_{\bullet} \bullet^{\bullet}\bullet\frac{\frac{\bullet\bullet}{\bullet}\bullet\frac{\bullet\bullet\bullet 1}{\bullet\bullet 1}\bullet\bullet\frac{\bullet\bullet}{\bullet\bullet\bullet}\bullet}{\bullet\bullet\bullet\bullet}$$

$$= \frac{\bullet\bullet}{\bullet\bullet\bullet\bullet} \sum_{\bullet} \bullet(\bullet\bullet 1) + 1\bullet\bullet\frac{\bullet\bullet\bullet 1}{\bullet\bullet 1}\bullet\bullet\frac{\bullet\bullet}{\bullet\bullet\bullet}\bullet$$

$$= \frac{\bullet\bullet}{\bullet\bullet\bullet\bullet}\bullet(\bullet\bullet\ \bullet\bullet) \sum_{\bullet} \bullet\frac{\bullet\bullet\bullet\bullet}{\bullet\bullet\bullet}\bullet\bullet\frac{\bullet\bullet}{\bullet\bullet\bullet}\bullet + \sum_{\bullet} \bullet\frac{\bullet\bullet\bullet\bullet}{\bullet\bullet\bullet}\bullet\bullet\frac{\bullet\bullet}{\bullet\bullet\bullet}\bullet\bullet$$

$$= \frac{\bullet\bullet}{\bullet\bullet\bullet\bullet}\bullet(\bullet\bullet\ \bullet\bullet)\bullet\frac{\bullet\bullet + \bullet\bullet\ \bullet\bullet}{\bullet\bullet\bullet + \bullet\bullet\bullet}\bullet + \bullet\frac{\bullet\bullet + \bullet\bullet\ \bullet\bullet}{\bullet\bullet\bullet + \bullet\bullet\bullet}\bullet\bullet$$

= •• / •••• •(•• • •)• ••• / ••• • + • ••• / ••• ••

= •• [(•• • •)(• • •) + (• • •)] / • • •

Entonces,

•˙(•) = •••˙• • ••(•)•˙

= •• [(•• • •)(• • •) + (• • •)] / • • • • (••)˙

= ••• • ••• / •••

Ejemplo 5.4:

En una empresa dedicada a la fabricación de llantas se dispone de un departamento de control de calidad en el cual se encuentran muestras de productos conformes y no conformes, se tienen 10 muestras de productos conforme y 7 de productos no conformes, si no se saben cuales son los conformes y los no conformes, y desea seleccionar al azar dos muestra de estos productos, suponiendo que al extraer una de las dos muestras miramos que es un producto no conforme, cual es la probabilidad de que la otra muestra también sea un producto no conforme?

Sea A el suceso que al seleccionar dos muestras al menos una sea un producto no conforme, y B el suceso de que al seleccionar dos muestras las dos sean productos no conformes?

Sea A el suceso que al seleccionar dos muestras al menos una sea un producto no conforme, y B el suceso de que al seleccionar dos muestras las dos sean productos no conformes

La probabilidad del suceso A es igual de que los dos productos sean no conformes, o que uno sea no conforme y el otro sea conforme, y se halla mediante la suma de dos probabilidades ajustables a la distribución hipergeométrica, la probabilidad de A será la siguiente:

$$\frac{\binom{7}{1}\binom{10}{1}}{\binom{17}{2}} + \frac{\binom{7}{2}\binom{10}{0}}{\binom{17}{2}} = 0.514 + 0.154 = 0.669$$

Y la probabilidad de B se halla como el cociente de los posibles casos favorables entre casos posibles, de la siguiente forma:

$$\frac{\binom{7}{2}}{\binom{17}{2}} = 0.154$$

Para hallar la probabilidad de que al conocer que una de las llantas seleccionadas sea un producto no conforme, la otra también sea un producto no conforme se realiza el cociente de la probabilidad B/A ya que se tiene una intersección entre los sucesos de A con B.

Entonces $p(B|A)=\frac{P(A\cap B)}{P(A)} = \frac{0.154}{0.669} = 0.2307$

5.1.1.7 Aproximación entre Distribuciones Discretas

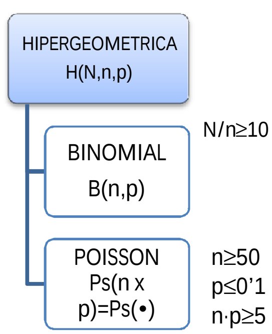
Durante el estudio que hemos hecho de las principales distribuciones discretas, hemos visto que éstas sólo se diferencian en pequeños matices. Por ejemplo, la distribución de Poisson es la misma que la Binomial pero cuando n es grande y p pequeña; la distribución Hipergeométrica también es la misma que la Binomial pero en el primer caso el tamaño de la población es finito y en el segundo es infinito.

Así pues, es evidente que éstas distribuciones pueden cierto parecido en el caso de que los parámetros tomen valores límites. En la figura siguiente. Se han representado los valores que deben tener dichos parámetros para poder realizar una buena aproximación.

Los valores que se dan en esta figura no son valores estándares aunque sí que son suficientes para hacer una buena aproximación. Dependiendo de la publicación éstos valores cambian:

Algunos autores recomiendan que sea $\lambda \geq 18$.

Figura 5.1 Aproximación de Distribuciones Discretas



Fuente: Los Autores

5.2 DISTRIBUCIÓN DE PROBABILIDAD PARA UNA CONTINUA

La Distribución de probabilidad •(•) proporciona el valor de probabilidad particular asociados a la función entre dos valores de la variable [•, •], bajo el área de la curva.

5.2.1 MODELOS DE PROBABILIDAD DE VARIABLES ALEATORIAS CONTINUAS

Dentro de los modelos de probabilidad, correspondiente a variables aleatorias continua, las de mayor aplicación se tienen: Normal Estándar, Uniforme Continua, Exponencial y Triangular. Además, los modelos de probabilidad correspondiente a variables aleatoria continuos se consideran los modelos normales estandarizados.

5.2.1.1 Distribución Normal

La distribución normal fue reconocida por primera vez por el francés Abraham de Moivre (1667-1754). Posteriormente, Carl Friedrich Gauss (1777-1855) elaboró desarrollos más profundos y formuló la ecuación de la curva; de ahí que también se la conozca, más comúnmente, como la "campana de Gauss". La distribución de una variable normal está completamente determinada por dos parámetros, su media y su desviación estándar (• y •).
Una variable aleatoria continua • sigue una distribución normal de media • y desviación típica •, si:

1. El recorrido de la variable • es toda la recta del conjunto de números reales.
2. La función de densidad es de la siguiente forma:

$$\bullet(\bullet) = \frac{1}{\bullet\bullet\sqrt{2\bullet}}\bullet^{(\bullet\bullet\bullet)^{\bullet}/\bullet\bullet^{\bullet}}$$

Hay muchas variables continuas cuya función de densidad tiene forma de campana. La distribución normal describe fenómenos en cuyo resultado final interviene gran número de factores independientes entre sí.
Ejemplo de esto es:

- Caracteres morfológicos de individuos (personas, animales y plantas) con características similares, ejemplo: Estatura, peso, etc.
- Caracteres fisiológicos como el efecto de un mismo fármaco
- Caracteres sociológicos

Las principales características de la función de densidad de la distribución normal son las siguientes:

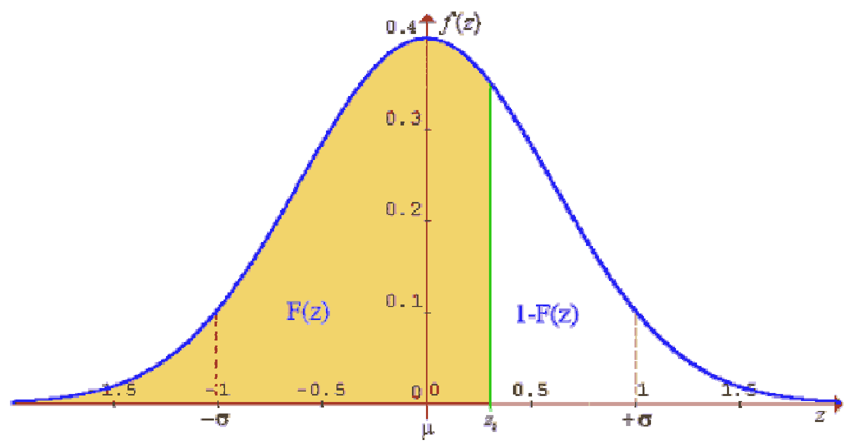
- Es simétrica respecto de la recta $\bullet = \bullet$, pues:

  $$\bullet(\bullet + \bullet_{\bullet}) = \bullet(\bullet\ \bullet\ \bullet_{\bullet}), \qquad \bullet\bullet_{\bullet}\ \bullet\ \bullet$$

  A causa de la simetría de la distribución normal de probabilidad, la mediana y la moda de la distribución también se hallan en el centro, por tanto en una curva normal, la media, la mediana y la moda poseen el mismo valor.
- Posee un máximo en el punto de abscisa $\bullet = \bullet$, y no tiene mínimos, por consiguiente es unimodal.
- Tiene dos puntos de inflexión en los puntos de abscisa $\bullet_{\bullet} = \bullet + \bullet$ y $\bullet_{\bullet} = \bullet\ \bullet\ \bullet$ por lo cual las dos colas (extremos) de una distribución normal de probabilidad se extienden de manera indefinida y nunca tocan el eje horizontal
- El eje de abscisas es una asíntota de la curva.
- Presenta forma de campana
- La media de una población distribuida normalmente se encuentra en el centro de su curva normal.

Para expresar que una variable continua •, tiene una distribución normal de media • y desviación típica •. Se escribe • (•, •)

Figura 5.2. Gráfica de la distribución normal



Fuente: tomado de http://personal5.iddeo.es/ztt/Tem/t21_distribucion_normal.htm

Ejemplo 5.5:

La media de los pesos de 500 estudiantes de un colegio es 70 kg y la desviación típica 3 kg. Suponiendo que los pesos se distribuyen normalmente, hallar cuántos estudiantes pesan:

1. Entre 60 kg y 65 kg.

$$\bullet(60 < \bullet < 75) = \bullet\bullet\frac{60 \bullet 70}{3} < \bullet < \frac{75 \bullet 70}{3}\bullet$$

$$\bullet(\bullet 3.33 < \bullet < 1.67) = \bullet(\bullet \bullet 1.67) \bullet [1 \bullet \bullet(\bullet \bullet 3.33]$$

$$0.9525 \bullet (1 \bullet 0.9996) = 0.9521(500)$$

2.Más de 90 kg.

$$\bullet(\bullet > 90) = \bullet\ \bullet\frac{90 \bullet 70}{3}\bullet$$

$$1 \bullet \bullet(\bullet < 6.67) = 1 \bullet 1 = 0{,}500 = 0$$

5.2.1.2 Distribución Normal Estándar $\bullet(\bullet,\bullet)$

La distribución normal estándar es aquella que corresponde a una distribución de media 0 y varianza 1. La función de densidad está definida por la expresión:

$$\bullet(\bullet) = \frac{1}{\sqrt{\bullet 2 \bullet}}\bullet^{\bullet\bullet\bullet\bullet\bullet\bullet\bullet};\quad \bullet\bullet < \bullet < \bullet$$

A partir de una variable $\bullet$ que tenga una distribución normal, se puede obtener otra característica z con una distribución normal estándar, a partir de la siguiente fórmula:

$$\bullet = \frac{\bullet\ \bullet\ \bullet}{\bullet}$$

Para determinar la probabilidad de una variable con distribución normal estándar, existen tablas.

5.2.1.3 Distribución Uniforme Continua

En la distribución de probabilidad uniforme, una variable $\bullet$, es aquella cuyos valores tienen la misma probabilidad.

Se dice que una variable aleatoria continua x tiene una distribución uniforme en el intervalo $[\bullet, \bullet]$ si la función de densidad de probabilidad es:

$$\bullet(\bullet) = \bullet\begin{cases} \dfrac{1}{\bullet\ \bullet\ \bullet} & \bullet\bullet\bullet\bullet\ \bullet\ \bullet\ \bullet\ \bullet\ \bullet \\ 0 & \bullet\bullet\bullet\bullet\ \bullet\bullet\ \bullet\bullet\bullet\bullet\bullet \end{cases}$$

La función de distribución en el caso continuo entre • y • es:

$$\bullet(\bullet) = \bullet \begin{cases} 0 & \bullet\bullet\bullet\bullet\ \bullet < \bullet \\ \frac{\bullet\ \bullet\ \bullet}{\bullet\ \bullet\ \bullet} & \bullet\bullet\bullet\bullet\ \bullet\ \bullet\ \bullet < \bullet\bullet \\ 1 & \bullet\bullet\bullet\bullet\ \bullet\ \bullet \end{cases}$$

Su media estadística es

$$\bullet(\bullet) = \frac{(\bullet + \bullet)}{2}$$

$$\bullet(\bullet) = \bullet_{\bullet\bullet}^{\bullet} \bullet\bullet(\bullet)\bullet\bullet = \bullet_{\bullet}^{\bullet} \bullet \frac{1}{\bullet\ \bullet\ \bullet}\bullet\bullet = \frac{1}{\bullet\ \bullet\ \bullet}\bullet\frac{\bullet^{\bullet}}{2}\bullet_{\bullet}^{\bullet} = \frac{\bullet + \bullet}{2}$$

y su varianza es:

$$\bullet^{\bullet}(\bullet) = \frac{(\bullet + \bullet)^{\bullet}}{12}$$

$$\bullet(\bullet^{\bullet}) = \bullet_{\bullet\bullet}^{\bullet} \bullet^{\bullet}\bullet(\bullet)\bullet\bullet = \bullet_{\bullet}^{\bullet}\bullet^{\bullet}\frac{1}{\bullet\ \bullet\ \bullet}\bullet\bullet = \frac{1}{\bullet\ \bullet\ \bullet}\bullet\frac{\bullet^{\bullet}}{3}\bullet_{\bullet}^{\bullet}$$

$$= \frac{1}{3}\frac{\bullet^{\bullet}\ \bullet\ \bullet^{\bullet}}{\bullet\ \bullet\ \bullet} = \frac{1}{3}(\bullet^{\bullet} + \bullet\bullet + \bullet^{\bullet})$$

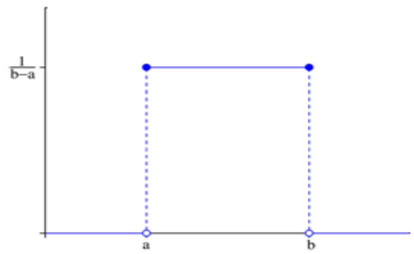
Entonces:

$$\bullet^{\bullet}(\bullet) = \bullet(\bullet^{\bullet}) \bullet [\bullet(\bullet)^{\bullet}] = \frac{1}{3}(\bullet^{\bullet} + \bullet\bullet + \bullet^{\bullet}) \bullet \bullet\frac{\bullet + \bullet}{2}\bullet^{\bullet}$$

$$= \frac{(\bullet^{\bullet} + \bullet\bullet + \bullet^{\bullet})}{3} \bullet \frac{(\bullet^{\bullet} + 2\bullet\bullet + \bullet^{\bullet})}{4}$$

$$= \frac{\bullet^{\bullet}\ \bullet\ 2\bullet\bullet + \bullet^{\bullet}}{12} = \frac{(\bullet\ \bullet\ \bullet)^{\bullet}}{12}$$

Figura 5.2 Grafica de Distribución Uniforme



Fuente: Los Autores

Ejemplo 5.6:

Una variable de comportamiento uniforme, presenta una media de 100 y una varianza de 6. Calcular que probabilidad que existe cuando • sea menor que 98.

En primer lugar se debe calcular • y •

Primero se sabe que

$$\bullet = \frac{\bullet + \bullet}{2} = 100$$

$$\bullet^{\bullet} = \frac{(\bullet \ \bullet \ \bullet)^{\bullet}}{12} = 6$$

$$\bullet = 97$$

$$\bullet = 103$$

La función de densidad de • es

$$\bullet(\bullet) = \bullet \begin{matrix} 0 \ \bullet\bullet \ \bullet \ \bullet \ 97 \\ \frac{1}{6} \ \bullet\bullet \ 97 \ \bullet \ \bullet \ \bullet \ 103^{\bullet} \\ 0 \ \bullet\bullet \ \bullet \ > 103 \end{matrix}$$

Entonces tenemos que $\bullet(\bullet \ \bullet \ 97) = \bullet_{\bullet \ \bullet}^{\bullet\bullet} \bullet(\bullet)\bullet\bullet = \bullet_{\bullet\bullet}^{\bullet\bullet} \frac{\bullet}{\bullet} \bullet\bullet = \frac{\bullet}{\bullet}$

5.2.1.4 Distribución Exponencial

La distribución exponencial es una distribución de probabilidad continua con un parámetro $\bullet = 0$ cuya función de densidad es:

$$\bullet(\bullet) = \bullet \begin{matrix} \bullet\bullet^{\bullet\bullet\bullet} & \bullet\bullet\bullet\bullet\ \bullet\ \bullet\ 0 \\ 0 & \bullet\bullet\ \bullet\bullet\bullet\bullet\ \bullet\bullet\bullet\bullet \end{matrix} \bullet$$

Su función de distribución es:

$$\bullet(\bullet) = \bullet(\bullet\ \bullet\ \bullet)\bullet \begin{matrix} 0 & \bullet\bullet\bullet\bullet\ \bullet < 0 \\ 1\ \bullet\ \bullet^{\bullet\bullet\bullet} & \bullet\bullet\bullet\bullet\bullet\ \bullet\ \bullet\ 0 \end{matrix} \bullet$$

El valor esperado y la varianza de una variable aleatoria $\bullet$ con distribución exponencial son

$$\bullet[\bullet] = \frac{1}{\bullet}$$

Y este se halla en primer lugar de su función característica

$$\bullet\bullet(\bullet) = \bullet \int_{\bullet}^{\bullet} \bullet^{\bullet\bullet\bullet}\bullet\bullet^{\bullet\bullet\bullet}\bullet\bullet = \bullet \int_{\bullet}^{\bullet} \bullet\bullet^{(\bullet\bullet\bullet\bullet)\bullet}\bullet\ \bullet = \bullet \frac{\bullet}{\bullet\bullet\ \bullet\ \bullet}$$

$$\bullet\bullet\bullet(\bullet) = \frac{\bullet\bullet}{(\bullet\bullet\ \bullet\ \bullet)^{\bullet}}$$

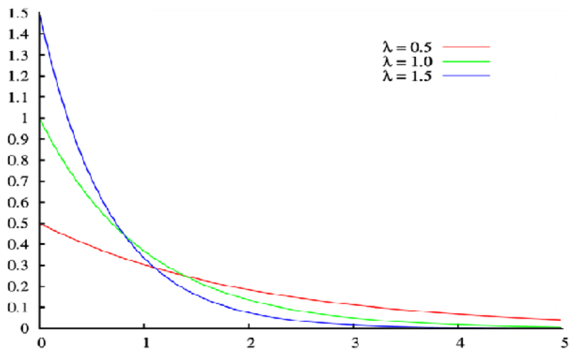
$$\bullet[\bullet] = \frac{\bullet\bullet^{\bullet}(0)}{\bullet} = \frac{1}{\bullet}$$

La derivada se haya realizando segunda derivada

$$\bullet^{\bullet} = \frac{\bullet}{\bullet^{\bullet}}$$

$$\bullet\bullet^{\bullet\bullet}(\bullet) = \frac{\bullet 2\bullet\bullet^{\bullet}}{(\bullet\bullet\ \bullet\ \bullet)^{\bullet}} = \frac{1}{\bullet^{\bullet}}$$

Figura 5.3 Grafica de Distribución Exponencial



Fuente: Los Autores

Ejemplo 5.7

En un aeropuerto se realizo un estudio de la llegada de los aviones de manera aleatoria e independiente y arrojo como resultado 180 aviones por hora. Utilizar la distribución exponencial, para encontrar la probabilidad de que un avión de manera aleatoria no llegue en medio minuto.

Teniendo en cuenta que X es la variable tiempo que transcurre entre dos llegadas consecutivas, se transforma las unidades de horas a minutos

$$\lambda = \frac{180}{60\ min} = 3 \qquad / min$$

Entonces tenemos que:

$$P(X > 0.5) = 1 - P(X < 0.5) = 1 - F(0.5)$$
$$= 1 - (1 - e^{-3 \cdot 0.5}) = e^{-1.5} = 0.22313$$

5.2.1.5 Distribución Triangular

Esta distribución tiene 3 parámetros, a (límite inferior de la variable); b (el modo) y c (límite superior de la variable).

Tabla 5.2 Distribución de Probabilidad Triangular

Triangular	TR (a, b, c)
Función de densidad	$\bullet(\bullet) = \frac{\bullet(\bullet\bullet\bullet)}{(\bullet\bullet\bullet)(\bullet\bullet\bullet)}$ si • • • • • $\bullet(\bullet) = \frac{\bullet(\bullet\bullet\bullet)}{(\bullet\bullet\bullet)(\bullet\bullet\bullet)}$ si • • • • •
Distribución acumulada	$\bullet(\bullet) = \frac{(\bullet\bullet\bullet)^{\bullet}}{(\bullet\bullet\bullet)(\bullet\bullet\bullet)}$ si • • • • • $\bullet(\bullet) = 1 \bullet \frac{(\bullet\bullet\bullet)^{\bullet}}{(\bullet\bullet\bullet)(\bullet\bullet\bullet)}$ si • • • • •
Parámetros	Parámetro de localización: • Parámetro de escala: •
Rango	a, b
Media	$\frac{\bullet + \bullet + \bullet}{3}$
Varianza	$\frac{\bullet^{\bullet} + \bullet^{\bullet} + \bullet^{\bullet} + \bullet\bullet \bullet \bullet\bullet \bullet \bullet\bullet}{18}$

Fuente: Adaptado y Tomado de “Estadística y Probabilidad de Ronald Walpol”

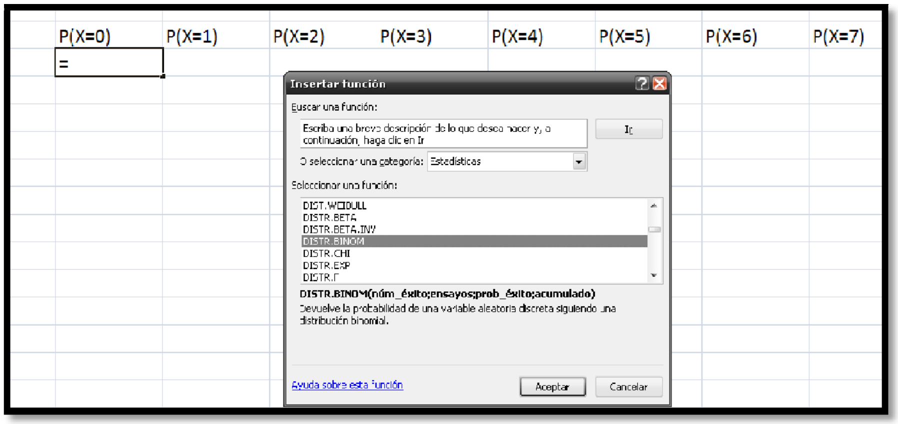
Se denomina triangular porque la función de densidad tiene una forma triangular, que viene definida en la tabla de anterior

Una distribución de probabilidad se denomina triangular cuando viene definida por dos parámetros, que representan el valor mínimo y el valor máximo de la variable. En este caso el triángulo es equilátero. Se denomina triangular (triangular general), cuando viene dada por tres parámetros, que representan el valor mínimo y el valor máximo de la variable, y el valor del punto en el que el triángulo toma su altura máxima. En este caso el triángulo no es necesariamente equilátero.

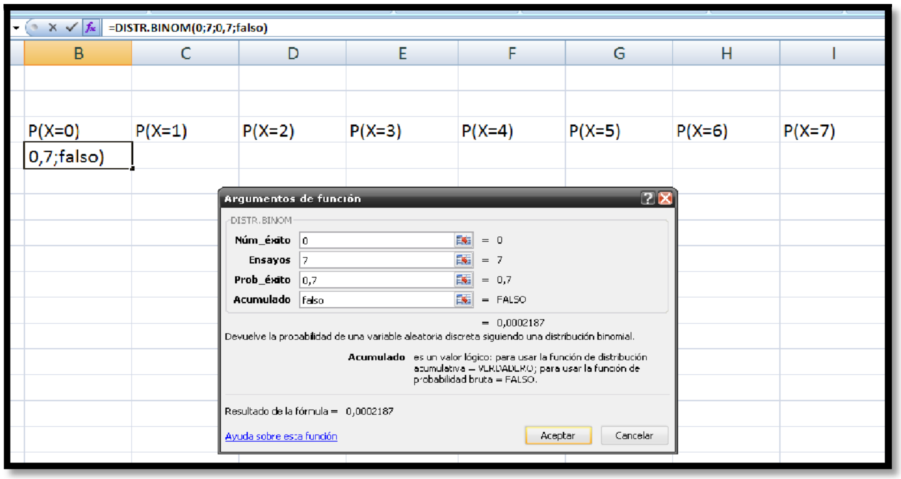
5.3 APLICACIÓN EN EXCEL

DISTRIBUCION BINOMIAL

Ejemplo 5.2:

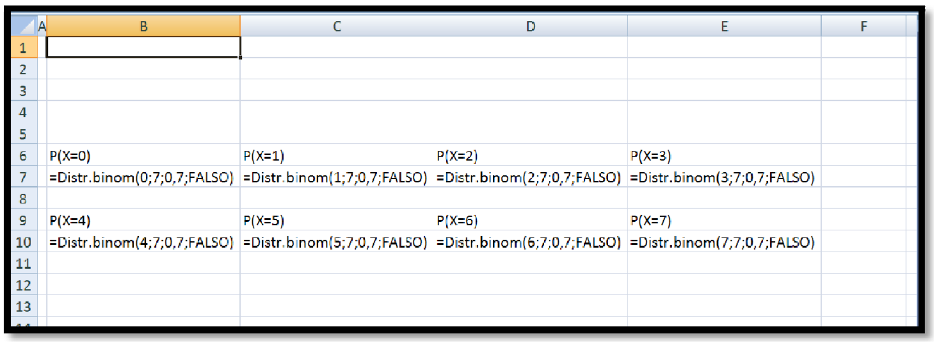


Primero se busca el tipo de distribución a utilizar, en esta ocasión se utiliza distribución binomial.



Se coloca la forma de distribución binomial de la forma =Dist.binom(0;7;0.7;Falso); el Falso es que no es acumulativo

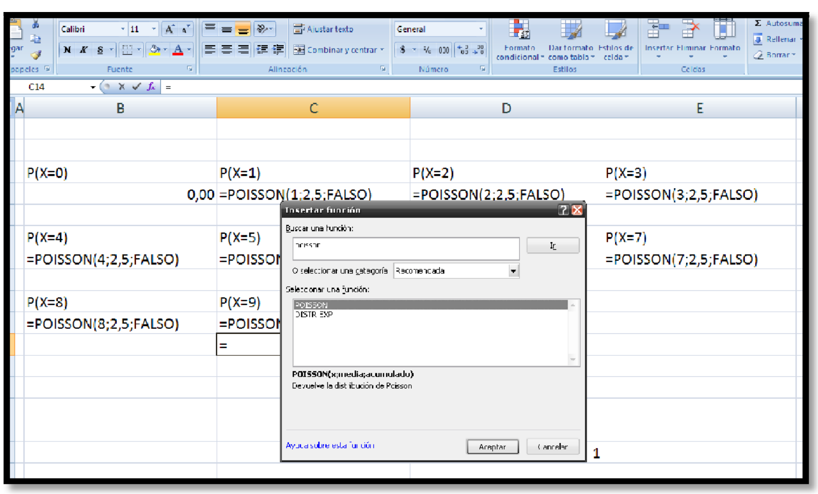
De igual forma se emplea para cada una de las *X* que se desean hallar;



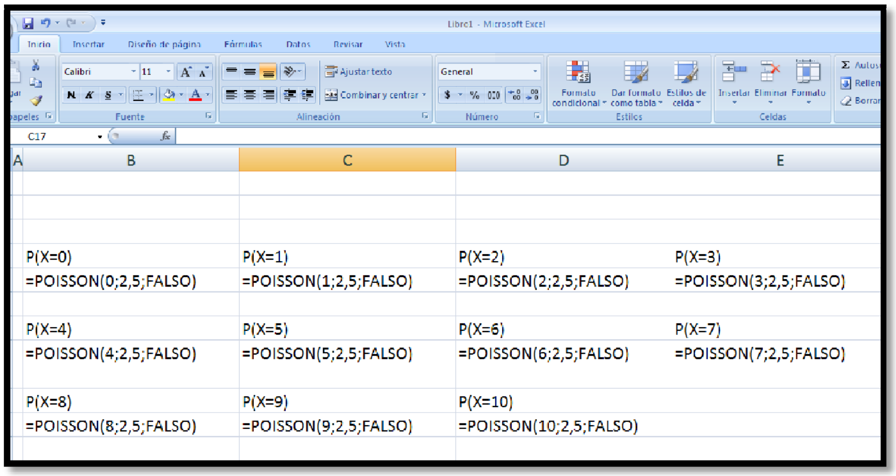
Y se hallan cada uno de los resultados

DISTRIBUCION DE POISSON

Ejemplo 5.3.



Primero se busca el tipo de distribución a utilizar, en esta ocasión se utiliza la función de distribución a aplicar

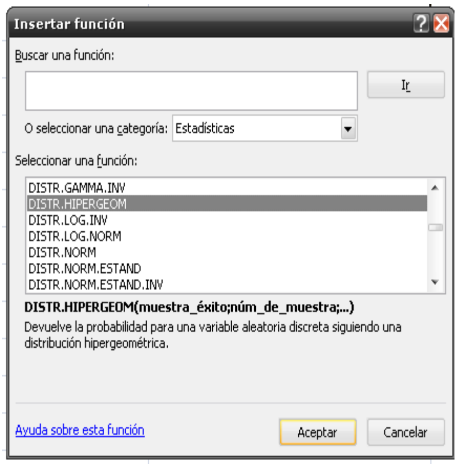


B	C	D	E
P(X=0)	P(X=1)	P(X=2)	P(X=3)
=POISSON(0;2,5;FALSO)	=POISSON(1;2,5;FALSO)	=POISSON(2;2,5;FALSO)	=POISSON(3;2,5;FALSO)
P(X=4)	P(X=5)	P(X=6)	P(X=7)
=POISSON(4;2,5;FALSO)	=POISSON(5;2,5;FALSO)	=POISSON(6;2,5;FALSO)	=POISSON(7;2,5;FALSO)
P(X=8)	P(X=9)	P(X=10)	
=POISSON(8;2,5;FALSO)	=POISSON(9;2,5;FALSO)	=POISSON(10;2,5;FALSO)	

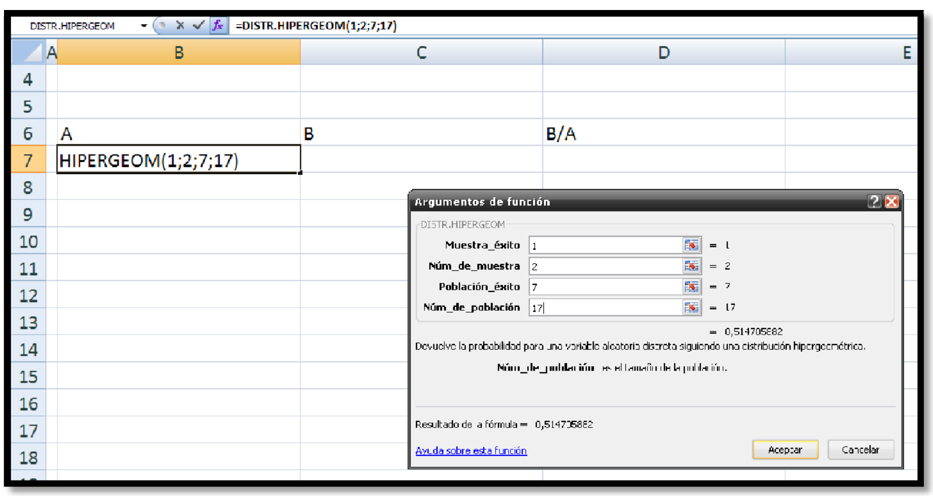
Se coloca la forma de distribución de poisson de la forma =POISSON (0; 2.5; Falso); el Falso es que no es acumulativo
De igual forma se emplea para cada una de las *X* que se desean hallar;

DISTRIBUCION HIPERGEOMETRICA

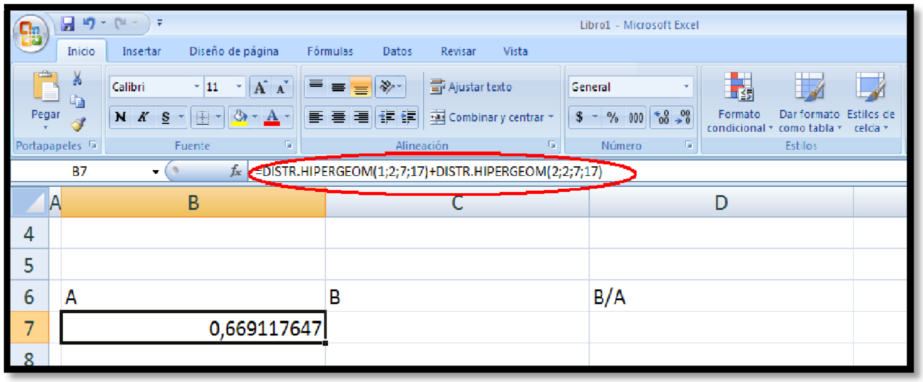
Ejemplo 5.4:
Primero se halla la función de distribución hipergeometrica para el suceso A, se busca en las funciones de Excel, en la parte de estadística



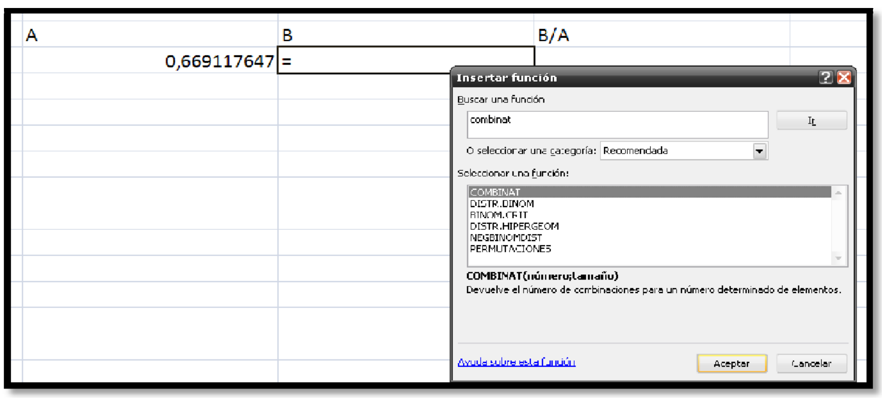
Segundo se insertan los datos allí solicitados



Esto es solo para cuando de la muestra una de las llantas es un producto no conforme, para cuando se halla la otra parte de A la que determina el caso cuando los dos productos son no conformes, se tiene que realizar una suma de las distribuciones hipergeometricas

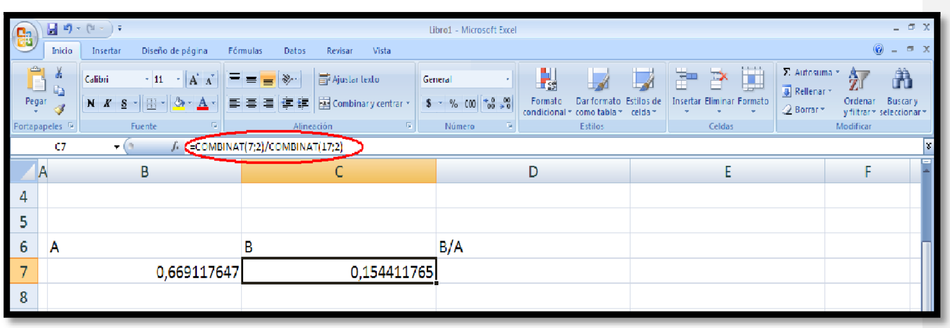


Para el suceso B se halla la combinatoria que las dos selecciones sean de productos no conformes.



De la forma =Combinat(7,2)/Combinat(17,2)

Y por último se halla el cociente entre B y A



EJERCICIOS PROPUESTOS

CAPÍTULO N° 5

EJERCICIOS PROPUESTOS

1. En un equipo de fútbol se desea alcanzar un total de un gol por la cantidad de jugadores que se encuentran en la cancha (11 jugadores), los cuales todos tienen un mismo ritmo de trabajo, la probabilidad de que un jugador no haga un gol es de 12/15. Determinar la probabilidad de que en un juego los jugadores
    1. Realicen goles todos los jugadores(11 Jugadores)
    2. Al menos 9 jugadores hagan goles
    3. Solo 7 jugadores hagan goles
    4. Al menos 3 jugadores hagan goles
    5. Ningún Jugador Realice Goles

2. En la empresa "Singer" se producen 60 maquinas de coser por día, si se seleccionan aleatoriamente 15 maquinas y se someten a una prueba para encontrar posibles defectos en ella, si 4 de las 60 maquinas están defectuosos:

    a) Cual es la probabilidad que la muestra contenga al menos 2 maquinas defectuosas?

    b) Cual es la probabilidad que solo haya 1 maquina defectuosas?

    c) Cual es la probabilidad que no existan maquinas defectuosas?

3. Se desea saber cuántos vacas en total se tienen en una finca, por esto se marcan 500 vacas y luego se dejan solas, días después se toman 100 vacas y se nota que 35 de ellas se encuentran marcadas. Se desea saber cuánto es el estimador máximo del tamaño de la población (N)?

4. En una sala de urgencias ingresan 225 personas enfermas. La probabilidad de que uno de los pacientes lleguen con problemas cardio-respiratorios es de 1/45. ¿Cuál es la probabilidad que en urgencias estén más de 3

personas con este tipo de problemas?, Cual es la probabilidad que en urgencias estén menos de 2 pacientes con este tipo de problemas?

5. En una tienda de ropa ingresan 55 compradores por hora.
    a) Cual es la probabilidad que no ingresen personas en un espacio de 15 minutos?
    b) Cual es la probabilidad que en 10 minutos ingresen 40 compradores?

6. En el departamento de control de calidad de una empresa de cementos se inspecciona el producto terminado. La proporción de productos defectuosos es de 0.017.
    a) Cual es la probabilidad de que la trigésima segunda unidad sea la quinta que se encuentra defectuosa?
    b) Cual es la probabilidad si se toma una muestra de 102 productos se encuentre solamente 1 con defecto?

7. En un casino existen diferente tipos de juegos como son los de poker, maquinas traga monedas, caballos, black jack y ruleta; la ruleta consta de 38 puestos que se conforman de la siguiente forma: 18 son rojos y negros, 2 son verdes.

    Si Y es una variable aleatoria que determina cual es el numero de veces que es necesario para girar la ruleta para obtener en la primera una posición en verde. Hallar la función de densidad Y?

8. Fabio Pedraza es el director del departamento de control de calidad de una marca muy reconocida de computadores portátiles, el realiza la revisión diaria de los inversores de la pantalla de los computadores portátiles. En la revisión se extraen 15 inversores y son revisados con el fin de encontrar algún defecto en ellos; según las estadísticas a lo largo del tiempo solo el

20% se convierten en productos no conformes, De acuerdo a lo anterior, responda:

a) Cuál es la probabilidad de que la muestra contenga más de 4 inversores defectuosos.
b) Cuál es la probabilidad de que ningún inversor se convierta en defectuoso?

9. Una maestra de física se encuentra muy escandalizada ya que cada vez que realiza un quiz se encuentra con que 13 de sus estudiantes realizan fraude en él, La maestra le informa a la coordinadora académica de la situación y le advierte que si llega a encontrar 3 estudiantes con fraude le anula el examen a todos los estudiantes. Cual es la probabilidad de que la profesora le anule todo los quices a los estudiantes?

10. En un supermercado se están dando puntos por la compra de cierta cantidad de productos; X es una variable con una distribución normal • • •. El supermercado decide asignar el puntaje de cliente platino a los clientes que excedan • + •, a los que tengan puntaje entre • • • + • como cliente Gold, si el cliente es VIP es por que su puntaje se encuentra entre el • • • y •, si es Importantes es porque su puntaje se encuentra con un puntaje entre • • 2• y • • •, y por ultimo si es de la categoría de comprador básico es porque su puntuación se encuentra inferior a • • 2•.
    a) Cual es el porcentaje de compradores que se encuentran en las diferentes categorías; Platino, Gold, VIP, Importante y Básico.

11. considere una variable aleatoria x con una distribución normal y una desviación típica igual a 2. Si se consideran muestras de tamaño 25, que distribución sigue la variable aleatoria?

ESTUDIO DE CASOS.

CASO 1

Comercializadora el Punto Ltda., es un distribuidor de ropa deportiva bastante reconocido de la cuidad.

El punto Ltda. Tiene varios proveedores, actualmente ropa sport le suministra los uniformes a un costo de 20.000 cada uno.

Por política de la empresa, los uniformes que no llegan a ser vendidos, son devueltos al proveedor quienes pagan 8.000 pesos por cada uno. Esto se hace con el fin de bajar los niveles de inventario de la empresa.

Para realizar el pedido a los proveedores, la empresa analiza y hace estimaciones de las ventas futuras teniendo en cuenta las ventas de meses anteriores. De acuerdo al departamento de mercadeo, las ventas registradas son las siguientes:

Tabla 5.3 Unidades Vendidas.

Mes	Unid. vendidas	Mes	Unid. vendidas
Enero 08	215	Enero 09	230
Febrero 08	200	Febrero 09	206
Marzo 08	185	Marzo 09	175
Abril 08	120	Abril 09	143
Mayo 08	145	Mayo 09	169
Junio 08	176	Junio 09	198
Julio 08	180	Julio 09	156
Agosto 08	135	Agosto 09	176
Septiembre 08	189	Septiembre 09	170
Agosto 08	145	Agosto 09	150
Octubre 08	180	Octubre 09	193
Noviembre 08	210	Noviembre 09	203
Diciembre 08	250	Diciembre 09	260

Fuente: Los Autores.

El gerente de ventas asegura que las ventas se comportan de acuerdo a una distribución normal con media 190 y desviación estándar de 30 unidades y que el pedido se debe realizar teniendo en cuenta la relación critica entre el costo de tener inventarios excedentes y el costo de tener inventarios faltantes. Es decir: costo de faltante/ (costo de excedente + costo de faltante).

Es importante destacar que el proveedor solo suministra cantidades especificas de uniformes, es decir producen por lote, un lote está constituido por 50 uniformes. Así que la empresa solo podrá pedir 50, 100, 150, 200, 250 o 300 unidades.

Podría pedir más de 300 unidades pero el gerente de ventas analiza que no es conveniente, puesto que las ventas en ningún mes han superado esta cantidad.
Está usted de acuerdo con el gerente de ventas en la afirmación que este realiza respecto al tipo de distribución de los datos. Justifique su respuesta.
De acuerdo la información de las ventas anteriores de la empresa, cree usted conveniente que el gerente se base en la relación crítica.

Como interpreta usted la relación critica desde el punto de vista de la estadística.
Teniendo en cuenta la relación critica, cuál es la cantidad de uniformes que la empresa debería comprar a su proveedor en enero del año 2010.

CASO 2.

Logitrans S.A es una empresa de logística de transporte y almacenamiento de mercancía, esta empresa cuenta con 5 agencias alrededor de Colombia, y 2 de sus agencias se encuentra en zona franca para minimizar algunos gastos de aranceles; además, reciben en puntos de servicio la mercancía con 3 horas antes de salir a su destino (estos envíos no supera el peso de los 50 kg).

La empresa tiene un reconocido nombre, pero en sus inicios tuvo fallas en los tiempos de entrega, lo que hacía que los clientes perdieran credibilidad en la misma, lo cual ocasionaba la perdida de mucho dinero, por incumplimiento en las clausulas de los contratos que se establecían con cada uno de los clientes.

En la medida que la empresa iba mejorando utilizó la planificación como medio para determinar los retrasos

Después de realizar estudios de tiempos, métodos y otra serie de estudios donde intervinieron deferentes variables que intervenían en el problema de los atrasos, y con los datos obtenidos se llego a la conclusión que el comportamiento de la variable que determinaba los retrasos se encontraba bajo la siguiente función de densidad

$$\bullet(\bullet) = \bullet\begin{matrix} \dfrac{1}{1+\bullet^{\bullet}} & \bullet\bullet\ 0 < \bullet < 1\bullet \\ 0 & \bullet\bullet\ \bullet\bullet\bullet\bullet\ \bullet\bullet\bullet\bullet \end{matrix}$$

Teniendo en cuenta la anterior función de densidad cual es el retraso medio de un pedido y la desviación típica. Para el 95% de los casos.

Si a usted le piden asesoría en la empresa Logitrans S.A que recomendaciones le daría según el comportamiento de la variable.

BIBLIOGRAFÍA

- DEVORE, Jay L. Probabilidad y estadística para ingeniería y ciencias. Séptima edición. México D.F. Cengage Learning. 2008

- INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS.Y CERTIFICACIÓN Trabajos Escritos: Presentación y referencias Bibliográficas. NTC 1486. BOGOTÁ. ICONTEC 2008.

- INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS.Y CERTIFICACIÓN Trabajos Escritos: Presentación y referencias Bibliográficas. NTC 5613. BOGOTÁ. ICONTEC 2008.

- LLINÁS S., Humberto. ROJAS, A., Carlos. Estadística descriptiva y distribuciones de probabilidades. Barranquilla. Ediciones Uninorte. 2005.

- MARTÍNEZ Bencardino, Ciro. Estadística y muestreo. Doceava edición. Bogotá D.C. ECOE. 2005.

- MASON D, ROBERT. Estadística para administración y economía. Décima Edición. México D.F. Alfa omega, 2001.

- MONTGOMERY, Douglas C. Probabilidad y Estadística aplicadas a la ingeniería. Segunda Edición. México D.F. Limusa. 2005

- MONTGOMERY, Douglas C. Diseño y análisis de Experimentos. Segunda Edición. México D.F. Limusa. 2007.

- PÉREZ López, César. Estadística aplicada a través de Excel. Madrid. Pearson. 2002.

- PEREZ López, César. Estadística práctica con Statgraphics. Madrid. Prentice Hall. 2002

- WALPOLE, Ronald E., MYERS, Raymond H., MYERS, Sharon L., YE, Keying., Probabilidad y estadística para ingeniería y ciencias. Octava Edición. México D.F. Pearson. 2007.

- WAYNE, DANIEL. Estadística con aplicaciones a las ciencias sociales y a la educación. México D.F. Mc Graw - Hill. 1988