



# ESTADÍSTICA I



# Índice

Presentación

Red de contenidos

Sesiones de aprendizaje

<b>SEMANA 1</b>	: Conceptos básicos, Población, Muestra, Estadígrafo y parámetros. Variables, clasificación. El estudio estadístico. Etapas Organización y clasificación de datos: cuadros y gráficos estadísticos.	9
<b>SEMANA 2</b>	: Cuadros de distribución de frecuencias. Variable discreta y continua	21
<b>SEMANA 3</b>	: Medidas de tendencia central. Datos agrupados y no agrupados. Propiedades	35
<b>SEMANA 4</b>	: Medidas de Posición, Percentiles Medidas de Dispersión.	43
<b>SEMANA 5</b>	: Análisis combinatorio. Probabilidad. Definiciones generales	55
<b>SEMANA 6</b>	: Álgebra de eventos. Propiedades Probabilidad condicional. Regla del producto.	67
<b>SEMANA 7</b>	: Probabilidad Total. Regla de Bayes. Diagrama del Árbol. Independencia Estadística.	79
<b>SEMANA 8</b>	: Distribución de probabilidad discreta Binomial	89
<b>SEMANA 9</b>	: <b>Semana de Exámenes Parciales</b>	
<b>SEMANA 10</b>	: Distribución de probabilidad discreta Poisson.	95
<b>SEMANA 11</b>	: Distribución de probabilidad continua. Distribución Normal. Características. Uso de la tabla de distribución Normal Estándar.	101
<b>SEMANA 12</b>	: Distribución Normal. Teorema del límite Central. Aproximación de las distribuciones discretas a la normal.	113
<b>SEMANA 13</b>	: Distribución muestral de medias y proporciones	121
<b>SEMANA 14</b>	: Control de calidad y procesos estadísticos. Administración de la calidad total. Gráfica de control para las medias <b>X</b> . Gráfica de control para la variación de procesos <b>R</b>	129
<b>SEMANA 15</b>	: Detección de tendencias en una gráfica de control. Gráfica de control para porcentajes <b>p</b> . Gráfica de control para el número de defectos por pieza <b>c</b> .	141
<b>SEMANA 16</b>	: Análisis de Regresión Simple. Diagrama de dispersión. Recta de mínimos cuadrados. Correlación entre dos variables. Aplicación e interpretación.	151
<b>SEMANA 17</b>	: <b>Semana de Exámenes Finales</b>	



# **Presentación**

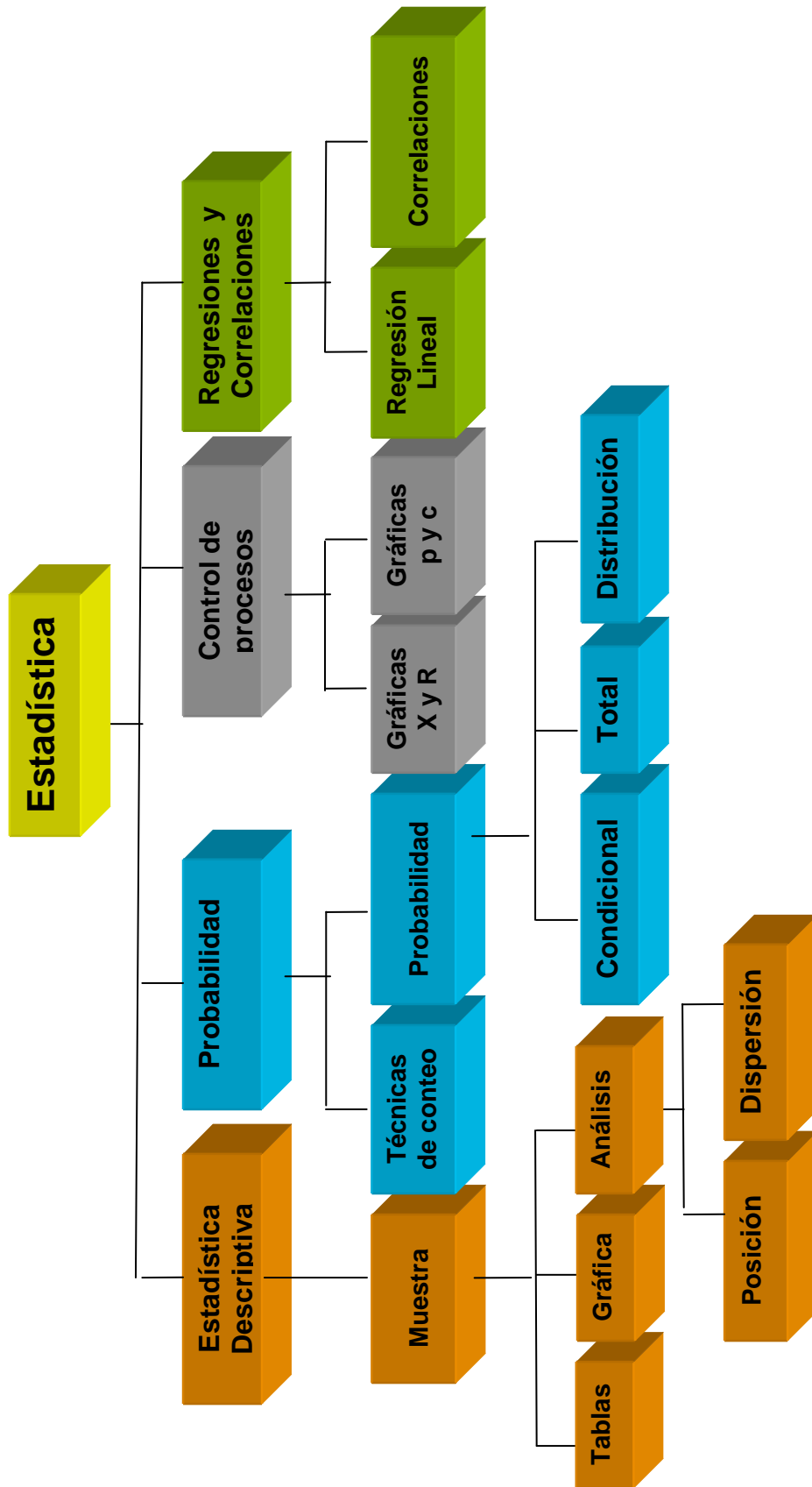
El avance de las comunicaciones y del software ha hecho que, en la actualidad, todo profesional esté en constante contacto con información estadística. Es más, muchas veces es necesario realizar alguna medición estadística para poder tener una idea acerca de la marcha de una empresa, o para tomar una decisión organizacional. El presente curso pretende ofrecer al futuro profesional, las herramientas necesarias para interpretar y evaluar información estadística, dándole además los fundamentos para realizarla.

En este curso se estudian los fundamentos teóricos de la estadística y la probabilidad, el análisis de cuadros estadísticos, la lógica en sus diferentes métodos de trabajo y los recursos para calcular y obtener las soluciones a los problemas. Además, se adquiere destreza en la interpretación y manejo de las definiciones, teoremas y fórmulas.

En una primera etapa, se desarrollará el marco teórico y práctico de la Estadística Descriptiva; la segunda etapa, comprenderá la aplicación básica de la Teoría de Probabilidades y, finalmente, en la tercera etapa se desarrollará la aplicación de las anteriores etapas, para un adecuado control de procesos estadísticos y proyecciones mediante regresiones.

Este curso es netamente práctico, por lo que en cada sesión se desarrollará la teoría necesaria en forma concreta, dándole mayor énfasis a la parte práctica y a la interpretación de resultados.

# Red de contenidos



La estadística como ciencia, nos proporciona un conjunto de métodos, técnicas o procedimientos para:



### 1. RECOPILAR



### 2. ORGANIZAR

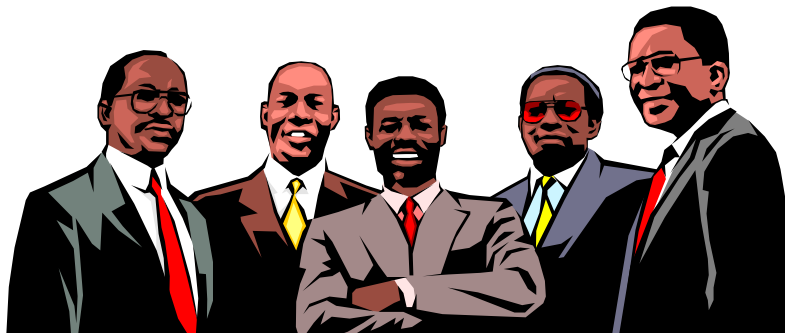


### 3. PRESENTAR



### 4. ANALIZAR

datos con el fin de describirlos o de realizar generalizaciones válidas con una medida de confiabilidad (probabilidad) para tomar decisiones.









# EL ESTUDIO ESTADÍSTICO

## TEMAS

- Conceptos básicos de la estadística.
- Variables clasificación
- El estudio estadístico. Etapas

## OBJETIVOS ESPECÍFICOS

- Distinguir los términos básicos de la estadística y las variables que intervienen en ella.
- Identificar las distintas etapas del estudio estadístico y qué se desarrolla en cada una de ellas.

## CONTENIDOS

- Definición de estadística.
- Definiciones básicas. Población, muestra, parámetro, estadígrafo.
- Variables Cualitativas y cuantitativas ( discretas y continuas)
- Objetivo.
- Etapas de la investigación estadística.

## ACTIVIDADES

- Identifican las diferentes definiciones existentes en estadística.
- Analizan las diferentes formas de una investigación estadística

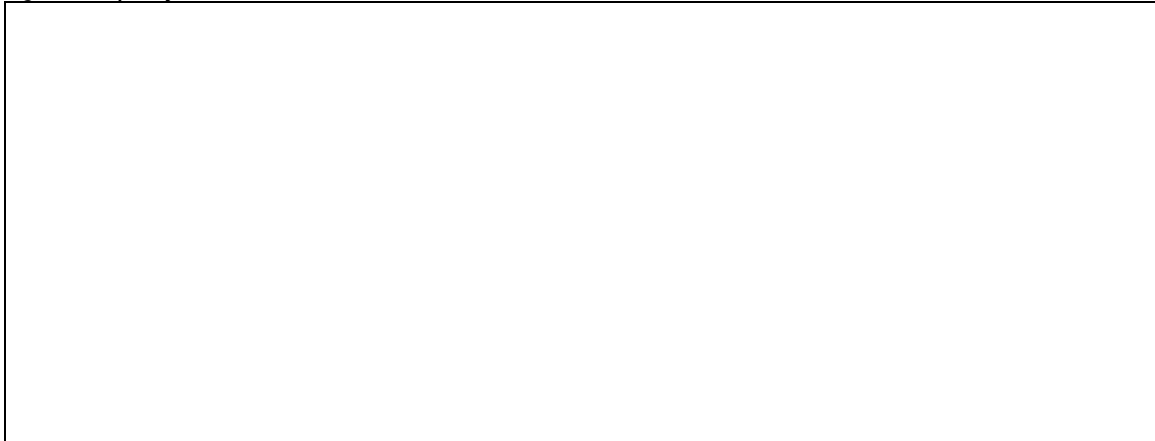
# ESTADÍSTICA Y PROBABILIDAD

## DEFINICIÓN E IMPORTANCIA

¿Qué es la Estadística?

Es una “Ciencia” que trata acerca de la recopilación, organización, presentación, análisis e interpretación de datos (numéricos), con el objeto de facilitar la toma de decisiones en cualquier campo.


¿Para qué y cuándo utilizamos la Estadística?



¿Cómo se divide la Estadística?

**La Estadística** se divide en:

1. **Estadística Descriptiva:** Es la que utiliza diversos procedimientos con el fin de recopilar, organizar, presentar y resumir un conjunto de datos.
2. **Estadística Inferencial o Inductiva:** Es la que utiliza diversos métodos para generalizar o inferir algo acerca de una “población” basados en resultados obtenidos en base a una “muestra”.



## Definiciones Básicas

**Población:** Es el conjunto conformado por todos los elementos materia de interés u objeto de estudio (personas, objetos o mediciones).

Ejemplos:

**Muestra:** Es una parte o subconjunto de la población materia de estudio.

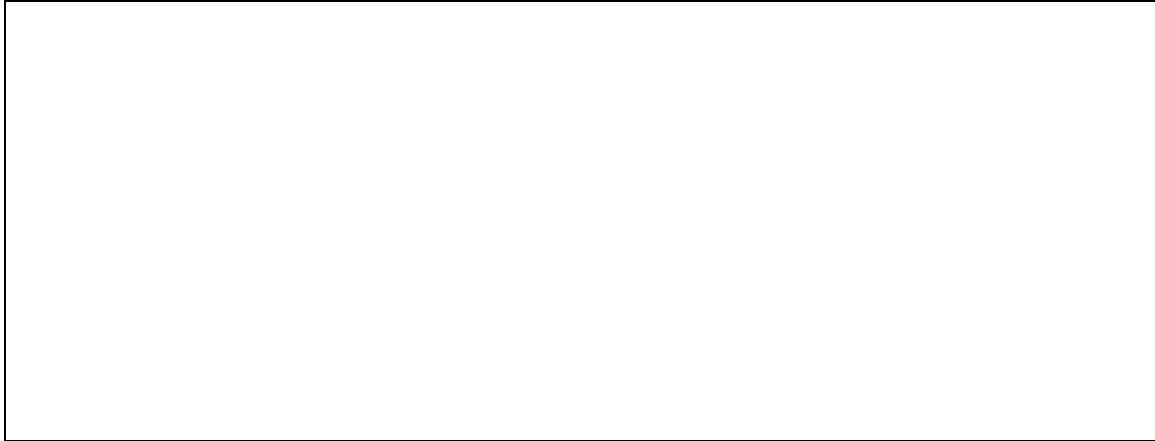
Ejemplos:

**Individuo o Unidad Elemental:** Es todo elemento que está afectado por la característica que se desea estudiar.

Ejemplos:

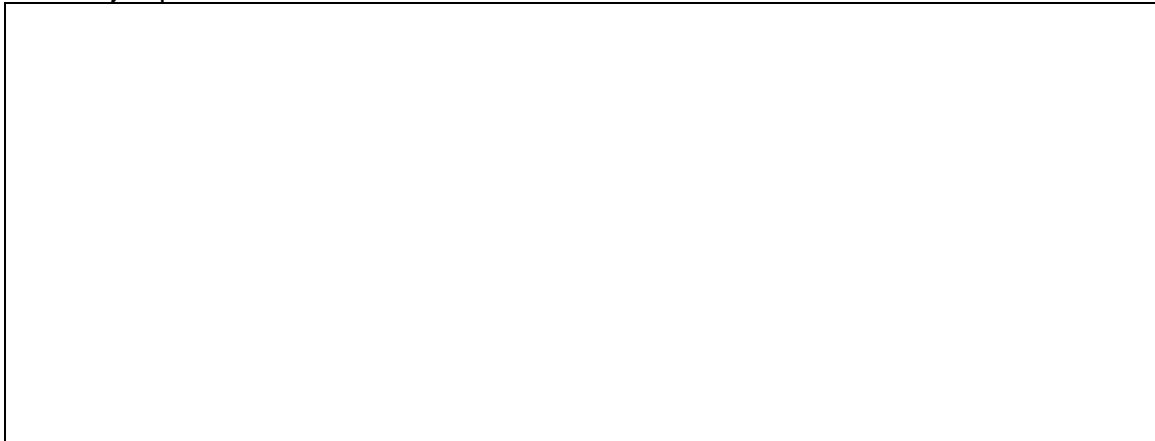
**Observación:** Es el dato o registro realizado, como consecuencia del resultado de la observación en la unidad elemental.

Ejemplos:



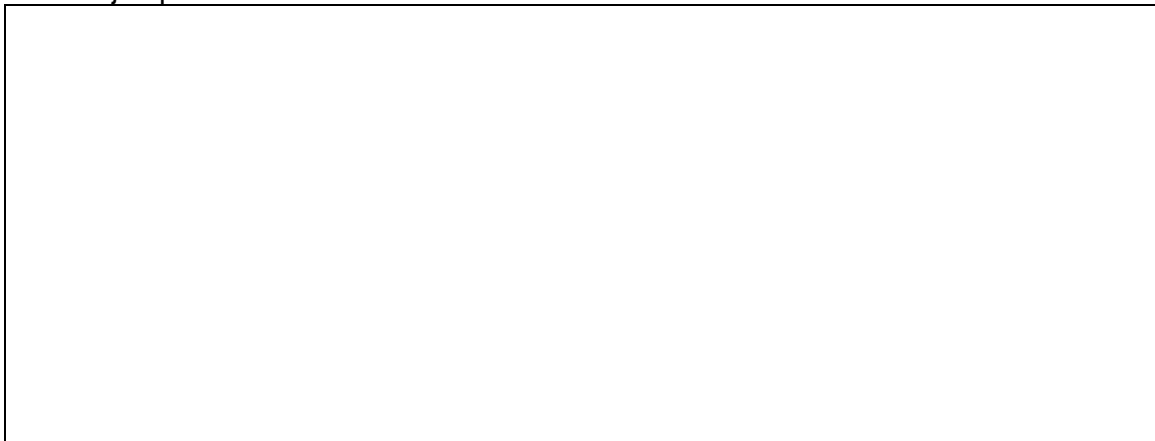
**Estadístico o Estadígrafo:** Es un valor representativo de la muestra

Ejemplos:



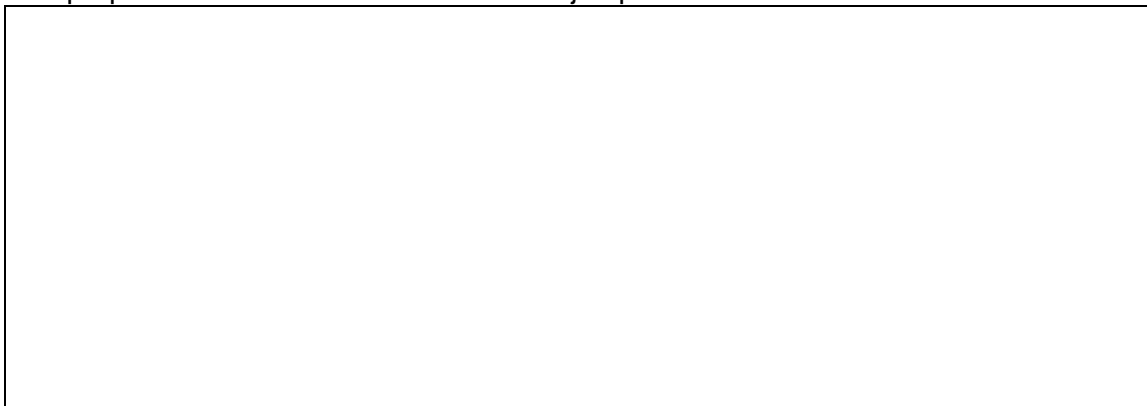
**Parámetro:** Es un valor representativo de la población

Ejemplos:



**Variable o carácter:** Es todo valor o característica que puede tomar valores diferentes cuando se observa a los individuos de una muestra o población. Pueden ser:

**a) Cualitativas:** Son aquellas que indican cualidad, atributo o categoría. No proporcionan valores numéricos. Por ejemplo:

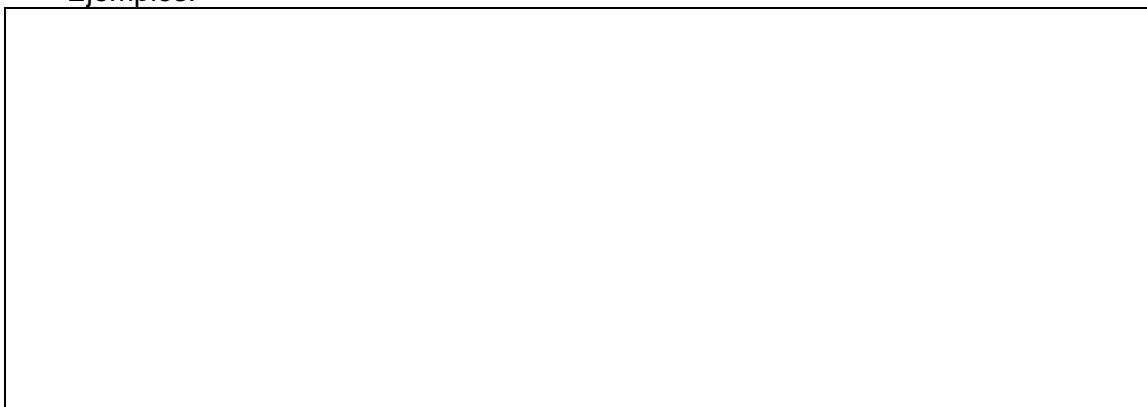


**b) Cuantitativas:** son aquellas que indican cantidad. Pueden ser:

**Discretas:** Cuando puede tomar un número finito de valores posibles. (Es numerable).

**Continuas:** cuando puede tomar un número infinito de valores Posibles. (No es numerable).

Ejemplos:



**Escala de medición:** Es un instrumento de medida, con el que se le asignan números a las unidades estadísticas para una variable definida.

Ejemplos:



# LA INVESTIGACIÓN ESTADÍSTICA

## Objetivo:

Obtener o estimar respuestas (con cierto grado de certeza) a determinadas interrogantes materia de estudio, utilizando métodos y procedimientos científicos.

Ejemplo:

Ganador de las elecciones municipales.

Determinación de la tasa de inflación para el próximo año.

## Etapas de La Investigación Estadística:

Planeamiento de la Investigación.

Recolección de los datos.

Organización y presentación de datos.

Análisis e interpretación de los resultados.

Generalización e Inferencia.

## Planeamiento de la Investigación

Se determina cuales son los objetivos de la investigación y los problemas que se presentan. Debe definirse qué información se debe recopilar, cómo y cuándo hacerlo. Se debe escoger una muestra representativa

## Recopilación de los datos

Se recogen los datos de acuerdo a los planes establecidos en la etapa anterior. El éxito de la investigación depende de la calidad de la información recopilada.

## Organización y Clasificación de los Datos

En esta etapa se clasifica a la información según sus características y se determina la mejor forma de presentarla mediante gráficos y cuadros.

## Interpretación de los Resultados

Se hallan estimadores y valores estadísticos que permitan una mejor interpretación de los mismos, según los criterios previamente establecidos.

## Generalización e Inferencia

Utilizando los métodos de Inferencia Estadística, las conclusiones de la investigación son generalizadas a la población de donde se obtuvo la información.

## ACTIVIDADES

1. Una empresa dedicada a la fabricación de conservas de pescado tiene planeado introducir al mercado conservas de trucha. Para ello le encargó a una empresa investigadora de mercado la realización de un estudio mediante el que le interesaba averiguar, entre otras cosas, la aceptación del nuevo producto y el precio que las personas estarían dispuestas a pagar. La encuesta fue realizada en Lima y se entrevistaron a 250 personas. De los encuestados, el 67% estarían dispuestos a consumir el nuevo producto. Además se concluyó que el precio del producto debería oscilar entre 1,50 y 2,5 soles. Determina:

- la población y la muestra,
- las variables y sus respectivos tipos,
- los estadísticos y los parámetros (si existen),
- el tipo de estimación que se utilizó.

2. Un empresario tiene la idea de implementar la venta de chicha morada en envase no retornable. Piensa que en un principio que debe analizar la posibilidad de lanzar su producto en lugares que sean cálidos durante gran parte del año. Para ello realiza un estudio de factibilidad en Piura e Iquitos. En Piura, el 90% de los 250 encuestados, está dispuesto a consumir el nuevo producto. En cambio, en Iquitos el 85% de los 300 encuestados muestra esta disposición. También obtuvo información acerca de la utilidad que conseguiría. En Piura, lograría un promedio de 1,5 soles de utilidad por producto, y en Iquitos, un promedio de 2 soles. Determina:

- la población y la muestra, las variables y sus respectivos tipos,
- los estadísticos y los parámetros (si existen).
- Si tuviese que elegir entre una de las dos ciudades para llevar a cabo su proyecto, ¿Cuál debería elegir? Justifique.

3. Suponga que la empresa XYZ realizó un estudio en la ciudad de Lima, sobre de las principales marcas de cerveza vendidas por el consorcio ABC, compuesto por los supermercados S y W. El estudio se realizó en diciembre de 2000, se entrevistaron a un total de 1200 consumidores elegidos al azar, 500 en el supermercado S y 700 en el supermercado W. Algunas de las informaciones se proporcionan a continuación:

- Las marcas de cerveza que se vendieron en el mes de agosto en estos supermercados fueron: Cristal, Pilsen, Cuzqueña, y Bremen.
  - El 30% de los consumidores son de condición media alta.
  - Tres personas que consumen no tienen preferencia por una marca en particular.
  - El 10% consume sólo la marca Bremen.
  - El 15% de los que prefieren Pilsen viven en el Callao o en zonas aledañas.
  - El 45% afirmó no tener problemas en consumir cualquier marca de cerveza.
  - Uno de los entrevistados pertenece a la clase media baja y prefiere la cerveza en lata.
  - Doscientos consumidores del mercado S y 80 del supermercado W afirmaron que la cerveza en lata presenta mejor sabor que en Six pack y que en envase retornable.
  - Uno de los objetivos del estudio era estimar la verdadera cantidad de dinero promedio que los consumidores gastan mensualmente en cerveza.
- De acuerdo al enunciado identifique:
- a) Población, muestra, variables y el tipo de variables.
  - b) A qué conceptos básicos corresponden las afirmaciones expresadas.



## Autoevaluación

1. Al medir cierta característica en una población, a una unidad estadística  $C_1$  se le asigna el valor 10 y a otra unidad estadística  $C_2$  se le asigna el valor 20. ¿Qué puede decir acerca de la característica de  $C_1$  y  $C_2$ ? si la escala usada es:
  - a) nominal
  - b) ordinal
  - c) de razón
2. Clasifique las variables e indique el tipo de escala en que se miden. Si las características observadas en un censo de alumnos de una determinada universidad particular son: Facultad, año de ingreso, tiempo de residencia en Lima, número de créditos acumulados, grado en la escala de pensiones, grado de instrucción del padre, número de hermanos, ingreso familiar mensual promedio.
3. Al investigar el nivel socioeconómico en las modalidades: bajo (B), medio (M), alto (A) de 20 familias, se obtuvo los siguientes datos: M, B, B, M, A, B, B, M, M, B, M, B, B, A, M, B, M, A, M, B. Construye la distribución de frecuencias y trazar su gráfica.
4. Al averiguar el número de hijos de 20 familias se obtuvieron los siguientes datos: 3, 2, 5, 0, 1, 3, 2, 1, 0, 1, 3, 4, 2, 4, 4, 3, 4, 3, 2, 3. Grafica la distribución de frecuencias relativas y frecuencias relativas acumuladas. Determine el porcentaje de familias que tiene dos hijos o más, pero menos de cuatro
5. La empresa XXX produce, entre otros productos, aceites y mantequilla para consumo humano. Dicha empresa encargó realizar un estudio sobre las preferencias de las personas de clase media de la ciudad de Lima acerca de dichos productos. Para el estudio se tomó una muestra aleatoria de 600 personas, y se obtuvo lo siguiente:

El 35% de los encuestados prefiere los productos “diet”  
El 55% de los encuestados prefería los aceites vegetales  
El 13% de los encuestados no supo diferenciar la mantequilla de la margarina  
Una persona dijo que no consumía ningún tipo de mantequilla.  
De acuerdo al enunciado identifique:

  - a) Población, muestra, variables y el tipo de variables.
  - b) Identifique a qué conceptos básicos corresponden las afirmaciones expresadas.
6. Una empresa desea probar la eficacia de un nuevo comercial de T.V. transmitido en la ciudad de “La Furia”. Como parte de la prueba, el comercial se pasa en el horario que se transmite el noticiero local. Dos días después se lleva una encuesta telefónica. De ella se obtuvo lo siguiente:

El 70% de los habitantes de la ciudad de “La Furia” sintoniza la estación de T.V. para ver el noticiero local.  
El 15% de las personas que vieron el comercial lo recordó inmediatamente.  
El 5% de las personas que recordaron el comercial compró el producto.  
Identifique:

  - a) Población y muestra
  - b) Las variables de estudio y su tipo.
  - c) Los estadísticos y parámetros (si existen)

7. Se realizó un estudio en el distrito XXX donde el principal objetivo era establecer la viabilidad de emprender un negocio de alquiler de computadoras. Era importante determinar si los habitantes del distrito poseían computadoras personales en casa o utilizaban las cabinas de Internet. El estudio arrojó, entre otros resultados, lo siguiente:  
El 80% de los 400 encuestados no poseen computadoras personales.  
El 20% afirmó que le interesaría alquilar por días alguna computadora personal.  
El 35% de los habitantes del distrito XXX tienen conocimientos básicos de Microsoft Office.  
Identifique:  
a) Población y muestra  
b) Las variables de estudio y su tipo.  
c) Los estadísticos y parámetros (si existen)
8. Una empresa realiza una encuesta sobre el uso de cajeros automáticos en un distrito de 50 000 personas. De las 1200 personas encuestadas se tiene la siguiente información:  
El 60% de los habitantes del distrito usa la red de cajeros Unibanca  
La edad promedio de los encuestados fue 38 años  
El 72% de los usuarios afirmó que únicamente los utiliza para retirar dinero  
El 5% de los encuestados afirmó haber sido asaltado al utilizar el cajero.  
Identifique:  
a) Población y muestra  
b) Las variables de estudio y su tipo.  
c) Los estadísticos y parámetros (si existen)
9. Se realizó un estudio para establecer las posibilidades que tenía el candidato X de ganar las elecciones municipales en el distrito A. Se encuestaron a 200 personas y se obtuvo lo siguiente:  
El 40% de los habitantes del distrito X no ha decidido aún su voto.  
El 23,5% indicó que votaría por el candidato X.  
El principal problema a resolver por el nuevo alcalde es el de la delincuencia, tal como lo indicaron el 75% de los encuestados.  
El 25% de los habitantes del distrito vive hace más de 5 años en dicha zona.  
Determinar:  
a) Población y muestra  
b) Las variables de estudio y su tipo.  
c) Los estadísticos y parámetros (si existen)

## Para recordar

- Una manera de averiguar cuál es la población en estudio es preguntarnos: ¿De quién obtenemos la información que necesitamos?
- Una misma variable estadística puede tener distintas clasificaciones. Establecer una de ellas va depender de las unidades que se utilicen.
- Los estadígrafos o parámetros provenientes de variables cualitativas se representan mediante proporciones (porcentajes); en cambio, los que provienen de variables cuantitativas se representan, por lo general, mediante promedios.
- Durante el, planeamiento de la investigación, se establece el tamaño de la muestra.
- Del cuidado que se tenga en la recolección de datos depende, en gran medida, el éxito de la investigación.





## **ORGANIZACIÓN Y CLASIFICACION DE LOS DATOS CUADROS DE DISTRIBUCIÓN DE FRECUENCIAS**

### **TEMAS**

- Organización y clasificación de datos:
- Cuadros estadísticos.
- Gráficos estadísticos.
- Cuadros de distribución de frecuencias.
- Variable discreta.
- Variable continua

### **OBJETIVOS ESPECÍFICOS**

- Presentar información en cuadros y gráficos estadísticos.
- Agrupar tablas de distribución de frecuencia un conjunto de datos discretos y continuos.

### **CONTENIDOS**

- Tipo de datos.
- Cuadros estadísticos.
- Graficas estadísticas
- Cuadros de distribución de frecuencias.
- Gráficos de distribución de frecuencias

### **ACTIVIDADES**

- Resuelven problemas y presentan información en cuadros y gráficos.
- Desarrollan cuadros y gráficos estadísticos.

## ORGANIZACIÓN Y CLASIFICACIÓN DE DATOS

Una vez recolectados, los datos se pueden clasificar mediante:

- Cuadros estadísticos
- Gráficos Estadísticos

### Cuadros de distribución de frecuencias

El tipo de cuadro o gráfico a utilizarse dependerá de la naturaleza de los datos

Tipos de Datos o Caracteres

Los datos o caracteres (parejas) pueden ser:

- Cualitativos
- Cuantitativos.

Mixtos, cuando un carácter es cualitativo y el otro cuantitativo.

### Cuadros Estadísticos

Se utilizan principalmente cuando los datos son cualitativos o uno de los datos es cualitativo y el otro es cuantitativo.

### CANTIDAD DE ALUMNOS POR CARRERA Y SEXO EN EL INSTITUTO ABC DURANTE 2005 II



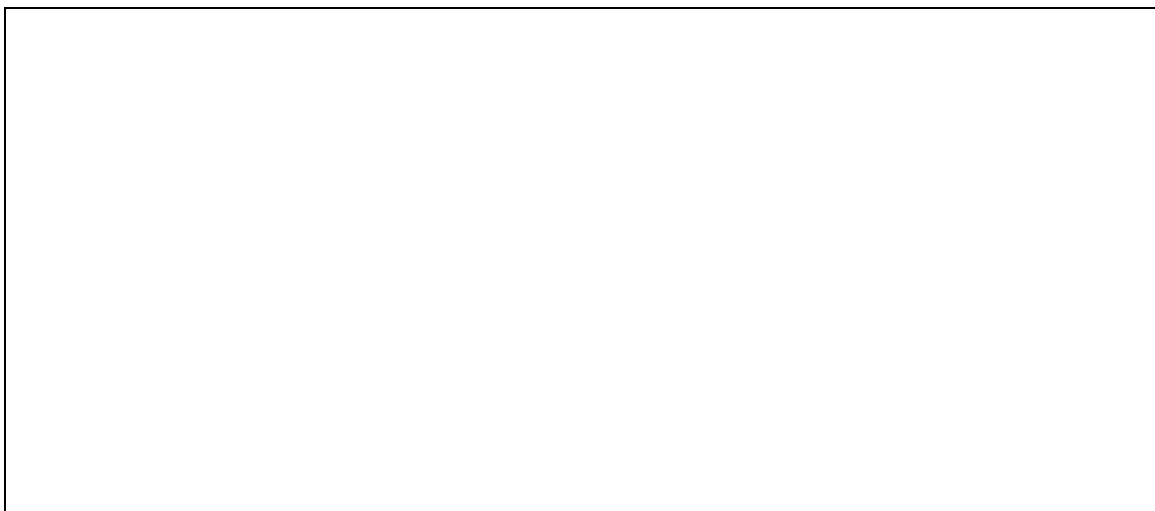
### GRÁFICOS ESTADÍSTICOS

El tipo de gráfico a utilizarse varía según la intención o los resultados obtenidos luego de la recolección de datos.

#### 1. Gráficos de Barras Simples

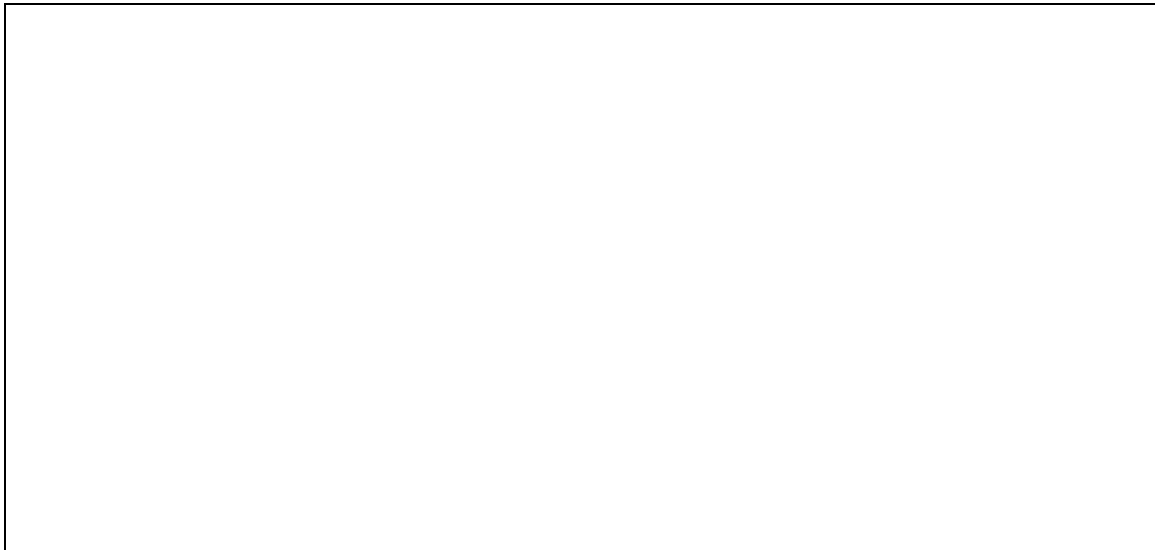
Cantidad de alumnos de sexo masculino matriculados el 2005-II por ciclo



**2. Gráficos de Barras Agrupadas****Cantidad de alumnos según sexo matriculados el 2005 II****3. Gráficos de Barras de Partes Componentes****Cantidad de alumnos según sexo matriculados el 2005 II por ciclo****4. Gráficos Circulares****Porcentaje de alumnos según sexo carreras matriculados el 2005 II**

## 5. Gráficos de Líneas

### Ventas realizadas entre 1995 y 2005



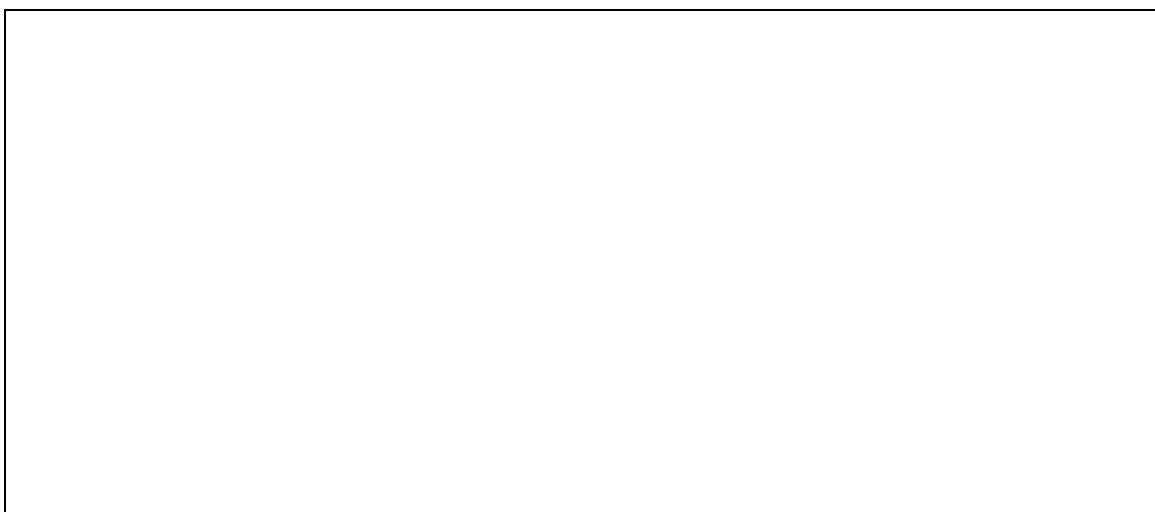
## Cuadros de Distribución de Frecuencias

Se utilizan principalmente cuando los datos son cuantitativos

### 1. Para variables cuantitativas discretas

A continuación se muestra el número de hijos de 36 familias de Lima:

2	3	4	5	1	2	3	2	1
0	2	1	5	3	1	2	3	2
2	4	3	5	2	0	2	1	3
1	1	4	2	3	4	5	1	0





## 2 Para variables cuantitativas continuas

2.1 Determinar el número de intervalos (K)

$$K \geq 1 + 3.3 \log(n)$$

2.2 Hallar el Rango (R) de los datos:

$$R = X_{\text{máx}} - X_{\text{mín}}$$

2.3 Determinar la amplitud (A) de los intervalos

$$A = R / K$$

2.4 Construir los intervalos: [ LI LS >

1er. Intervalo: L.I.:  $X_{\text{mín}}$  L.S.:  $X_{\text{mín}} + A$

2do. Intervalo: L.I.:  $X_{\text{mín}} + A$  L.S.:  $X_{\text{mín}} + 2A$

3er. Intervalo: L.I.:  $X_{\text{mín}} + 2A$  L.S.:  $X_{\text{mín}} + 3A$

4to. Intervalo: L.I.:  $X_{\text{mín}} + 3A$  L.S.:  $X_{\text{mín}} + 4A$

Y así sucesivamente hasta llegar al último intervalo.

K-ésimo Intervalo: L.I.:  $X_{\text{mín}} + (k-1)A$

L.S.:  $X_{\text{máx.}}$

2.5 Determinar la Marca de clase ( $m'_i$ ) de cada intervalo:

$$m'_i = (LI + LS)/2$$

La marca de clase es el valor representativo del intervalo (Valor medio).

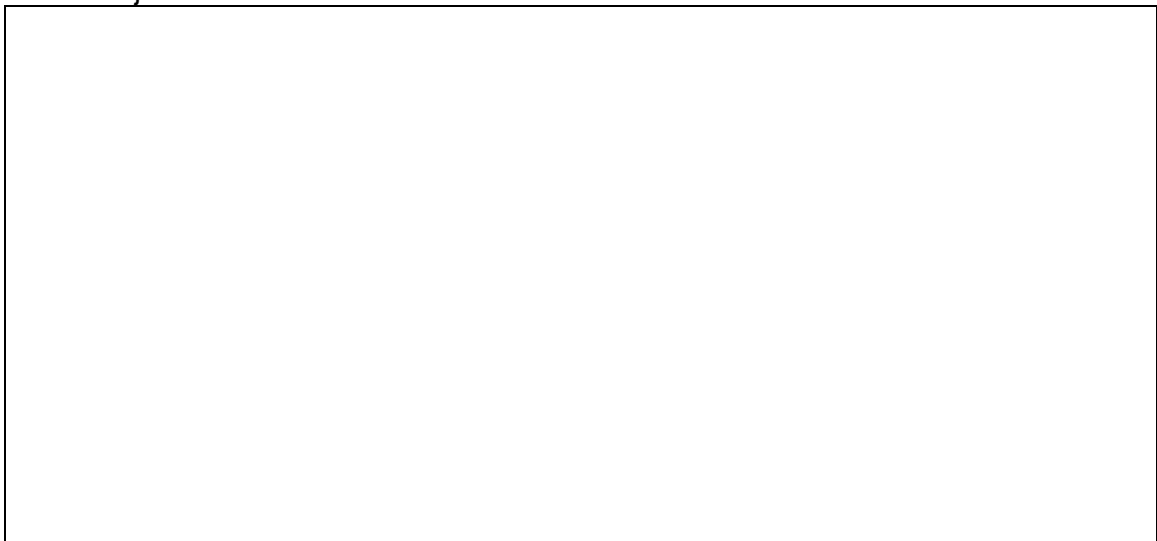
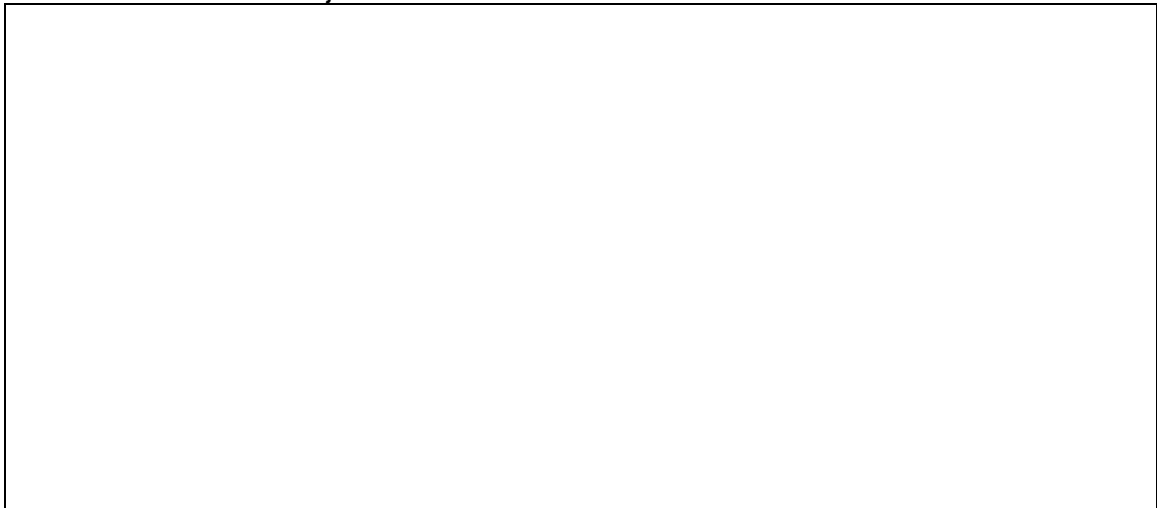
2.6 Halla las frecuencias absolutas ( $f_i$ ) de cada intervalo, contabilizando el número de datos de la muestra que pertenecen a cada intervalo.

2.7 Las frecuencias absolutas acumuladas ( $F_i$ ), así como las frecuencias relativas ( $h_i$  y  $H_i$ ) se hallan de la misma forma que para una distribución de frecuencias de variables cuantitativa discreta.

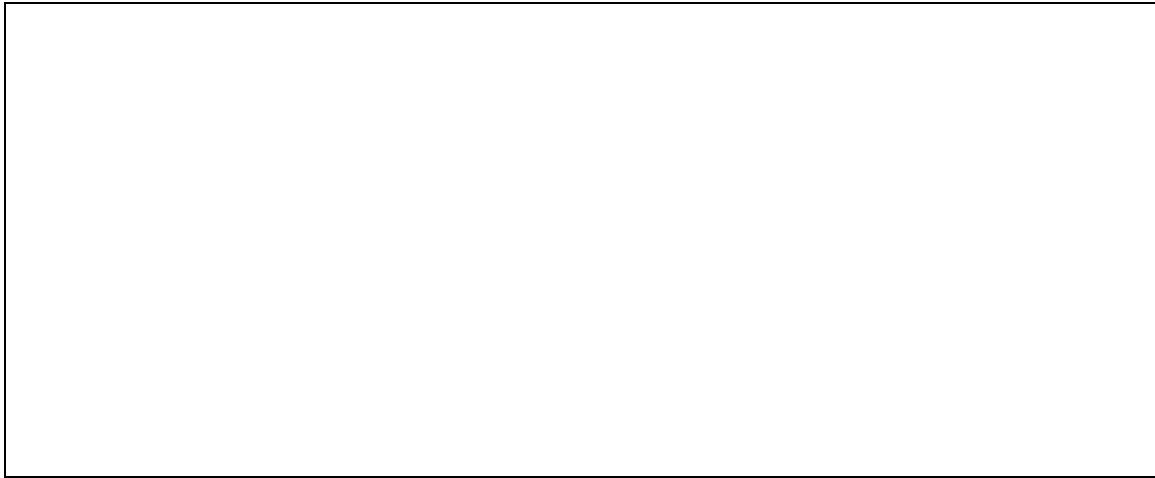
Ejemplo:

Los siguientes datos son los puntajes obtenidos por 50 estudiantes en un examen:  
Distribución de Puntajes

33	50	61	69	80	35	52	64	71	81
35	53	65	73	84	39	54	65	73	85
41	55	65	74	85	41	55	66	74	88
42	57	66	76	89	45	59	66	77	91
47	48	60	68	78	97	60	67	77	94



- 3.3 Ojiva:** Eje X: Límite superior de cada intervalo  
Eje Y: Frecuencia Absoluta (relativa) acumulada



## ACTIVIDADES

1. El siguiente conjunto de datos se refiere a los sueldos mensuales (en dólares) de 40 empleados de una empresa:

440	560	335	587	613	400	424	466	565	393
453	650	407	376	470	560	321	500	528	526
570	430	618	537	409	600	557	432	591	428
440	340	558	460	560	607	382	671	512	492

Construir una tabla de distribución de frecuencias con igual intervalos de clase, y elabore el histograma

2. A fin de decidir cuantos mostradores de servicio se necesitarán en tiendas que serán construidas en el futuro, una cadena de supermercados quiso obtener información acerca del tiempo (en minutos) requerido para atender a los clientes. Se registraron los siguientes tiempos:

3,6	1,3	0,8	1,3	1,1	0,8	1,0
0,3	0,9	0,7	3,1	2,2	1,6	1,9
0,4	5,1	1,8	0,3	1,1	0,6	0,7
1,1	1,9	2,1	0,3	0,8	1,0	1,4
1,1	0,5	1,2	0,6	0,8	1,7	0,2

2.1 Construya una tabla de distribución de frecuencias.

2.2 Interprete  $f_2$ ,  $X^3$  y  $H_4$ .

2.3

3. El tiempo en minutos que demoran 30 alumnos de la sección en terminar de contestar esta pregunta son:

3,36	4,28	2,32	3,12	3,86	2,92
2,08	3,26	2,42	4,00	2,12	2,68
2,18	4,28	2,22	3,96	3,36	2,12
4,14	2,02	2,82	3,96	2,42	4,24
3,22	4,02	2,68	3,86	2,92	3,56

3.1 Construya una tabla de distribución de frecuencias donde se clasifiquen estos datos, intérprete:  $X_4$ ,  $f_1$ ,  $F_2$ ,  $h_4$ ,  $H_3$

3.2 Construya el histograma, el polígono de frecuencia y la Ojiva

- 4 En un estudio sobre la opinión de los trabajadores de una empresa acerca de sus ingresos se ha obtenido los siguientes resultados:

Ingresos	Hombres	Mujeres
Muy inconforme	60	25
Inconforme	80	40
Resignado	95	95
Conforme	45	80

4.1 Establece un gráfico que permita comparar las opiniones de los hombres y mujeres acerca de sus ingresos.

4.2 ¿Qué grupo tiene un mayor porcentaje de personas que no están conformes con sus ingresos?

- 5 Una empresa que se dedica a preparar dietas realizó un muestreo a 50 empleados chequeando sus pesos, obteniéndose los siguientes resultados:

76,9	77,4	67,9	63,7	49,9	46,4	68,8	67,3	72,3	75,8
88,3	94,6	57,3	87,3	74,3	73,2	90,4	76,3	52,7	71,7
75,6	41,8	73,6	71,4	83,2	67,4	99,3	62,3	89,2	86,8
65,2	62,1	44,8	82,9	81,7	70,4	74,6	76,9	85,7	40,9
80,2	60,2	71,6	77,1	94,4	61,4	82,1	78,3	51,2	79,3

5.1 Elabore un cuadro de distribución de frecuencias.

5.2 Construya las gráficas correspondientes

## Autoevaluación

1. Algunas de las informaciones obtenidas por la empresa XYZ, se muestran a continuación:

Marca de Cerveza que compraron los consumidores de cada mercado

Marca de Cerveza	Número de consumidores	
	S	W
Cristal	250	200
Pilsen	90	180
Cuzqueña	150	240
Bremen	10	80
<b>Total</b>	500	700

Cantidad de dinero que se invierte mensualmente en Cerveza en cada Supermercado (en nuevos soles)

Supermercado S		Supermercado W	
Cantidad de dinero	Cantidad de consumidores	Cantidad de dinero	Cantidad de consumidores
[20, 40[	120	[40, 60[	280
[40, 60[	140	[60, 80[	196
[60, 80[	100	[80, 100[	126
[80, 100[	80	[100, 120[	56
[100, 120[	30	[120, 140[	21
[120, 140[	20	[140, 160[	14
[140, 160[	10	[160, 180[	7
<b>Total</b>	500	<b>Total</b>	700

- 1.1 Construya un solo sistema que permita comparar el consumo de cerveza de cada uno de los supermercados con la del consorcio. Comente y analice.
- 1.2 Establece un gráfico que permita comparar la distribución de gastos en cerveza de los consumidores de cada uno de los supermercados con el consorcio. Comente y analice.
- 1.3 Si el consorcio tiene en total 50 mil clientes habituales, estime la cantidad total de clientes que gastan en promedio más de 85 nuevos soles mensualmente.
- 2 A continuación se muestra el número de artículos defectuosos producidos durante los 30 días de un mes por una fábrica:

2	4	3	1	2	0	4	3	1	2
1	2	3	0	2	0	1	2	3	2
1	2	3	2	1	2	2	4	3	2

Construya el cuadro de distribución de frecuencias correspondiente.

3. Determina los intervalos de la distribución de frecuencias para cada uno de los siguientes casos:

- Datos siguientes,  $X_{\min} = 10$ ,  $X_{\max} = 50$ , y  $k = 8$  intervalos.
- Datos con dos decimales,  $X_{\min} = 2.55$ ,  $X_{\max} = 3.86$ , y  $k = 7$ .
- Datos con dos decimales,  $X_{\min} = 0.28$ ,  $X_{\max} = 0.65$ , y  $k = 6$ .

4. El peso en gramos de 30 objetos de un mismo tipo son los siguientes:

21,3	15,8	18,4	22,7	19,6	15,8	26,4	17,3	11,2	23,9
26,8	22,7	18,0	20,5	11,0	18,5	23,0	24,6	20,1	16,2
18,3	21,9	12,3	22,3	13,4	17,9	12,2	13,4	15,1	19,1

4.1 Construye una distribución de frecuencias de 6 intervalos de clase.

4.2 Calcula el porcentaje de objetos cuyo peso sea 21.5 gr o más.

4.3 Halla el peso debajo del cual se encuentran el 25% de los objetos.

5. En una compañía el sueldo mínimo de 200 empleados es de \$ 150. Si se sabe que 20 empleados ganan al menos 150\$, pero menos de \$ 180, 60 ganan menos de 210\$, 110 gana menos de \$240, 180 ganan menos de \$270 y el 10% restante de empleados ganan a lo más \$ 300; reconstruir la distribución y graficar su polígono de frecuencias.

6. El consumo mensual de agua (en metros cúbicos) de ochenta familias, se tabuló en una distribución de frecuencias simétricas de 7 intervalos de amplitud iguales a 3, siendo las marcas de clase del cuarto intervalo igual 19. Si las frecuencias del primer y tercer intervalo son iguales al 5% y 15% del total respectivamente, y si la quinta frecuencia porcentual acumulada es 85% del total.

6.1 Reconstruye la distribución.

6.2 ¿Cuántas familias consumen a lo mas 18.25 metros cúbicos al mes?

7. Los puntajes de una prueba de aptitud se tabularon en una distribución de frecuencias de 6 intervalos de igual amplitud.

Si se tienen marcas de clase:  $m_2 = 40$  y  $m_4 = 80$ ,

Frecuencias:  $h_1 = h_6$ ;  $h_3 = h_5$ ;  $h_4 = 0,25$ ;  $h_2 = h_4 - h_1$ ;  $h_3 = h_1 + 0,10$  y  $F_6 = 60$ . Completa la distribución de frecuencias absolutas y grafica el polígono.

8. El tiempo (en horas) de 120 familias que utilizan su computadora se tabularon en una distribución de frecuencias de 5 intervalos de amplitud iguales a 4, siendo: el tiempo mínimo de uso 2 horas, la primera y segunda frecuencias iguales al 10% y 15% del total de casos respectivamente. Si el 73,75% de las familias lo usaron menos de 17 horas y el 85% menos de 19 horas. Determine las frecuencias.
9. Las notas de un examen se tabularon en una distribución de frecuencias relativas de 3 intervalos de amplitud iguales a 5. Si la nota mínima es igual a 5, el 48% de las notas son menores que 12, y si el 80% de las notas son inferiores a 16, reconstruya la distribución de frecuencias.



10. Los datos que se muestran a continuación corresponden a las edades de 50 beneficiarios de un programa de asistencia social del gobierno:

81	53	67	60	80	64	56	54	91	61
66	88	67	65	97	72	74	65	73	69
43	54	76	70	86	68	82	75	79	60
41	87	76	97	70	45	60	45	65	56
92	72	82	80	52	65	50	58	70	76

10.1 Construir una tabla de distribución de frecuencias.

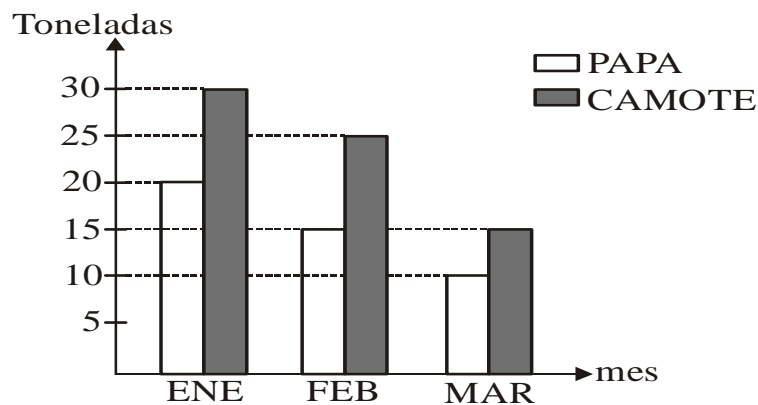
0.2 A partir de la tabla de distribución de frecuencias, determine el porcentaje de personas que tienen más de 72 años.

11. A continuación se muestran los tiempos (en segundos) que una máquina transportadora se demora en trasladar 40 paquetes:

44.0	56.8	46.5	48.7	36.3	42.5	32.4	46.6	35.5	39.3
45.3	56.6	43.7	37.6	42.3	45.6	33.5	47.8	32.8	42.6
57.0	43.9	36.8	45.7	34.9	48.6	35.7	43.2	39.1	38.8
44.0	33.4	35.8	46.0	36.8	32.7	38.2	37.1	45.2	39.2

Construya un cuadro de distribución de frecuencias de 5 intervalos de igual amplitud.

12. El gráfico muestra la producción (en ton.) de los tubérculos, en tres meses del año.



12.1 ¿En qué porcentaje descende la producción de camote entre febrero y marzo?

12.2 ¿En qué porcentaje difiere la producción de camote respecto a la de papa en el mes de marzo?

12.3 ¿Cuál es el promedio de la producción de papa en los tres meses?

## Para recordar

- Mostrar información en un gráfico es más ventajoso que hacerlo en un cuadro estadístico pues a simple vista podemos sacar conclusiones acerca de la información que estamos recibiendo.
- En la actualidad, gracias al distinto software existente, el cuadro estadístico ha pasado a ser una herramienta de referencia para elaborar gráficos.
- En la mayoría de los casos la información recolectada se puede representar en cualquier gráfico, sin embargo, si se desea mostrar la tendencia de la variable es recomendable utilizar el gráfico de líneas; si se quiere comparar varios totales es mejor utilizar un gráfico de barras; si se quiere mostrar como están distribuidas las partes componentes de un total se utiliza el gráfico circular; etc.
- Los cuadros de distribución de frecuencias tienen como principal función facilitar el cálculo de los estadísticos o parámetros adecuados.
- La Ojiva nos permite obtener valores por debajo del cuál se encuentra un porcentaje determinado de datos.
- El histograma y el polígono de frecuencias nos permiten tener una idea acerca de la naturaleza de la distribución de datos. En cierta forma nos permiten intuir, cuando la distribución lo permite, cuál es la probabilidad de que en la población se obtenga un rango de valores.



## **MEDIDAS DE TENDENCIA CENTRAL**

### **TEMAS**

- Medidas de tendencia central.
- Datos agrupados y no agrupados.
- Propiedades

### **OBJETIVOS ESPECÍFICOS**

- Calcular e interpretar los estadígrafos de tendencia central.

### **CONTENIDOS**

- Medidas de tendencia central: Media, Mediana, Moda  
Para datos no tabulados, tabulados discretos y continuos.

### **ACTIVIDADES**

- Interpretan los valores de tendencia central hallados

## MEDIDAS DE TENDENCIA CENTRAL

**Medidas de Tendencia:** Son aquellos valores que se toman como referencia para señalar el comportamiento de un conjunto de datos. Los más utilizados son: la media aritmética o promedio, la mediana, la moda y los percentiles. Estos últimos son denominados medidas de posición.

1. **Media Aritmética o promedio:** Se define como la suma de todas las observaciones (datos) dividida entre el número de observaciones. Es el valor representativo de los datos si es que entre las observaciones no hay valores extremos que influyen negativamente sucediendo lo mismo si los datos son muy dispersos.

Cálculo de la Media Aritmética

Para datos no agrupados	Para datos agrupados	
	Discretos	Continuos
$\bar{X} = \frac{\sum x_i}{n}$	$\bar{X} = \frac{\sum x_i \cdot f_i}{n}$ $\bar{X} = \sum x_i \cdot h_i$	$\bar{X} = \frac{\sum m'_i \cdot f_i}{n}$ $\bar{X} = \sum m'_i \cdot h_i$

Donde:

- $x_i$ : Valor observado
- $m'_i$ : Marca de clase
- $f_i$ : Frecuencia absoluta
- $h_i$ : Frecuencia relativa

2. **Mediana (Me):** Es el valor que divide a un conjunto de datos ordenados en dos partes iguales. Si el número de datos es impar la mediana es el valor central del conjunto de datos. Si el número de datos es par, la mediana es el promedio de los datos centrales.

Cálculo de la Mediana para datos agrupados:

$$Me = L_j + A_j \left( \frac{n/2 - F_{j-1}}{f_j} \right)$$

donde:

- $L_j$ : límite inferior del intervalo que contiene a la mediana.
- $f_j$ : frecuencia absoluta del intervalo que contiene a la mediana
- $F_{j-1}$ : frecuencia absoluta acumulada del intervalo anterior al que contiene a la mediana.
- $A_j$ : amplitud del intervalo que contiene a la mediana

3. **Moda (Mo):** Es el valor que más se repite (más frecuente. La desventaja de la moda radica en que en un conjunto de datos puede existir más de una distribución bimodal.

Para su cálculo se utiliza la siguiente fórmula (datos agrupados):

$$Mo = L_j + A_j \left( \frac{f_j - f_{j-1}}{(f_j - f_{j-1}) + (f_j - f_{j+1})} \right)$$

donde:     $L_j$ :        límite inferior del intervalo modal  
               $f_j$ :        frecuencia absoluta del intervalo modal  
               $f_{j-1}$ :      frecuencia absoluta del intervalo anterior al intervalo modal  
               $f_{j+1}$ :      frecuencia absoluta del intervalo siguiente al intervalo modal

## ACTIVIDADES

1. A continuación se muestran los sueldos de los 35 empleados de una empresa:

250	350	450	780	591	289	358
485	540	675	551	850	328	428
470	258	500	600	300	548	752
300	620	360	528	284	820	700
460	490	510	480	568	800	690

1.1  
Const  
ruya  
una  
tabla  
de  
distrib  
ución

de frecuencias.

1.2 Determine cuál es la medida de tendencia central que mejor representa a los datos. Justifique.

1.3 Calcule el valor por encima del cual se encuentra el 80% de los datos.

2. En una librería determinada que vende enciclopedias, se contabilizó la cantidad de ventas en los últimos 80 días, obteniéndose la siguiente información:

<b>Enciclopedias Vendidas</b>	0	1	2	3	4	5	6
<b>Cantidad de días</b>	7	11	18	20	19	1	5

2.1 Construya la tabla de frecuencias correspondiente

2.2 Interprete  $f_1$ ,  $F_4$ ,  $h_3$ ,  $H_2$

2.3 Calcule las medidas de tendencia central. Interpretelas

--

3. Los sueldos mensuales (en dólares) de 60 empleados de la empresa Pirámide S.A. son los siguientes:

440	560	335	587	613	400	424	466	585	393
453	650	407	376	470	560	321	500	528	526
570	430	618	537	409	600	557	432	591	428
440	340	558	460	560	607	382	671	512	492
450	530	501	471	660	470	364	634	580	450
574	500	462	380	518	480	625	507	645	382

3.1 Agrupe los datos en una tabla de distribución de frecuencias

3-2 Interprete  $f_1$ ,  $f_4$ ,  $h_2$ ,  $H_3$ .

3.3 Construya el Histograma correspondiente

3.4 Calcule las medidas de tendencia central. Interpretelas

--

4. La ojiva de una distribución de frecuencias está definida por los puntos: (100,0), (200, 5), (300, 13), (400, 28), (500, 48), (600, 60), (700, 70). La distribución se refiere a las ventas realizadas por los vendedores de una empresa que opera en todo el país durante el último mes.

4.1 Determine la mediana y la moda. Interprete

4.2 Halle el porcentaje de inversiones menores a 538,50 soles.

- 5 Un proveedor de servicios de Internet ha llevado a cabo un estudio para evaluar la frecuencia de uso de este servicio en las dos facultades que existen. Se realizó una encuesta a fines de octubre de 2005 en ambas facultades evaluándose el número de horas de uso del servicio de Internet y el gasto (nuevos soles) por el uso de dicho servicio. Se obtuvieron los siguientes resultados:

Minutos/día de Internet	Facultad A		Facultad B	
	Estudiantes	Gasto	Estudiantes	Gasto
Menos de 20	20	2,5	7	2,6
[ 20 40>	30	3,8	10	3,5
[ 40 60>	140	6,0	20	6,4
[ 60 80>	150	7,8	35	8,1
[ 80 100>	50	9,6	65	10,1
[100 120]	10	11,5	23	12,3
Total	400		160	

Efectúe un análisis comparativo sobre el tiempo de uso del servicio de Internet de los estudiantes de ambas facultades. Utilice las medidas estadísticas necesarias.



## Autoevaluación

1. Los costos de fabricación (en soles) de diez objetos son los siguientes: 9,35; 9,46; 9,20; 9,80; 9,77; 9,00; 9,99; 9,36; 9,50; 9,60. Si el precio de venta de cada objeto es 3 veces su costo de fabricación menos 5 soles, calcular la utilidad media por objeto.
2. En una evaluación, 5 alumnos tienen cada uno nota 12, y un alumno tiene 18. Si se indica como nota promedio 13, ¿qué nota promedio es?, ¿Es el promedio adecuado?, ¿Cuánto es el promedio adecuado?
3. De las edades de cuatro personas, se sabe que la media es igual a 24 años, la mediana es 23 y la moda es 22. Encuentre las edades de las cuatro personas.
4. De la curva de frecuencias de los sueldos de 30 empleados de una empresa, se sabe que  $Mo = \$200$ ,  $Me = \$220$ , y  $\bar{x} = \$250$ . Califique como verdadera o falsa las siguientes afirmaciones, justificando su respuesta:
  - 4.1 El sueldo más frecuente es de \$200 y más de la mitad de los empleados gana más de esa cantidad.
  - 4.2 Con una suma de \$3 300 se asegura el pago de la mitad de los empleados y con \$7 500 el de todos los empleados.
5. El sueldo promedio de 200 empleados de una empresa es S/400. Se proponen dos alternativas de aumento: a) S/. 75 a cada uno, b) 15% de su sueldo más 10 soles a cada uno. Si la empresa dispone a lo más de S/. 94 000 para pagar sueldos, ¿cuál alternativa es más conveniente?
6. Cuatro fábricas A, B, C y D, producen un mismo objeto. La fábrica B produce el doble de C, la D 10% menos que la C y la A al 60% menos que la B. Los costos de producción (en dólares) por unidad de estas fábricas son respectivamente: 0.2; 0.3, 0.2, y 0.5. Halle el precio medio de venta si se quiere ganar el 20% por unidad.
7. Las puntuaciones de un test aplicado a un grupo de estudiantes, se tabularon en una distribución de frecuencias de 6 intervalos de igual amplitud de manera que la marca de clase del segundo intervalo es 25 y el límite superior del quinto intervalo es 60. Si las frecuencias en porcentajes del primer al cuarto intervalo son respectivamente 15, 20, 35, 14 y si el 94% de las puntuaciones son menores que 60.
  - 7.1 Determine la distribución de frecuencias de las puntuaciones y graficar el polígono de frecuencias.
  - 7.2 ¿qué porcentajes de estudiantes tienen entre 38 y 53 puntos?
8. Los ingresos familiares, en decenas de dólares, se tabularon en una distribución de frecuencias simétricas de 6 intervalos, siendo las frecuencias:  $f_3 = 30$ ,  $f_3 = f_1 + 5$ , y  $F_6 = 150$ . Si el límite inferior del sexto intervalo igual a 60, y si el 75% de los ingresos son mayores de 43,5. Completar la distribución de frecuencias.
9. Al tabular las calificaciones de un examen se obtuvieron las siguientes notas: 07, 08, 09, 10, 11, 12, 13, 14, 15, 16, 17 y las frecuencias del número de alumnos respectivos: 1, 1, 1, 1, 1, 6, 8, 16, 18, 20, 2.
  - 9.1 ¿Cuánto es la media, la mediana y la moda de las notas?, ¿Qué valor escogería como el promedio?

9.2 ¿ Cuánto es la nota mínima para estar en el quinto superior?

10. La temperatura en los días de Abril fue la siguiente:

26,0	29,9	27,8	27,4	28,1	28,1
27,1	30,1	27,4	28,3	27,6	28,4
27,5	28,0	28,9	28,5	29,2	27,9
27,6	27,6	27,5	25,8	29,7	30,1
29,4	24,6	27,7	26,2	26,0	28,9

10.1 Construya una tabla de distribución de frecuencias donde se clasifiquen estos datos.

10.2 Interprete:  $X_4$ ,  $f_1$ ,  $F_2$ ,  $h_4$ ,  $H_3$

10.3 Elabore el histograma y las Ojivas

10.4 Calcule e interprete las medidas de posición.

## Para recordar

- El promedio es, por lo general, la medida que mejor representa los datos.
- Si los datos son muy dispersos o encontramos valores extremos es posible que el promedio no sea representativo de los mismos. En este caso es mejor utilizar la mediana.
- Cuando el cálculo de las medidas de tendencia central se hace en base a cuadros de distribución de frecuencias los resultados son aproximados.



SEMANA

4

## MEDIDAS DE POSICION

## MEDIDAS DE DISPERSION

### TEMAS

- Aplicaciones de las Medidas de Posición,
- Percentiles
- Medidas de Dispersión.

### OBJETIVOS ESPECÍFICOS

- Calcular e interpretar los estadígrafos de Posición.
- Calcular e interpretar los estadígrafos de Dispersión.

### CONTENIDOS

- Aplicaciones de las medidas de posición
- Percentiles
- Definición de variabilidad
- Varianza para datos no agrupados, agrupados discretos y continuos
- Coeficiente de variación
- 

### ACTIVIDADES

- Hallan el percentil del 25%, 50% y 75%
- Analizan los datos finales de las medidas de dispersión.

## PERCENTIL

Se denomina **percentil**  $k$  ( $k = 1, 2, 3, \dots, 99$ ) al valor numérico que denotaremos por  $P_k$ , que deja por debajo de sí el  $k$  por 100 de los datos ordenados. Así, el percentil 25,  $P_{25}$ , deja por debajo de sí el 25% de datos de la serie que se trate. El percentil 50,  $P_{50}$ , deja por debajo de sí el 50% de datos. El percentil 75,  $P_{75}$ , deja por debajo de sí el 75% de datos de la serie que se trate.

Observar que el percentil 50,  $P_{50}$ , es igual a la mediana.

**CUARTIL.** Se denomina cuartil a cada uno de los tres percentiles  $P_{25}$ ,  $P_{50}$ ,  $P_{75}$ . El cuartil uno,  $Q_1$ , es igual al percentil,  $P_{25}$ , el cuartil dos, o la mediana,  $Q_2$ , es igual al percentil  $P_{50}$ , y el cuartil tres,  $Q_3$ , es igual al percentil  $P_{75}$ .

El percentil  $P_k$ ,  $K = 1, 2, 3, \dots, 99$ , de  $n$  datos tabulados por intervalos, se deduce por interpolación, resultando:

$$P_k = L_{Pk} + \left[ \frac{n(k/100) - F_{i-1}}{f_i} \right] A$$

Donde:  $L_{Pk}$  es el límite inferior del intervalo que contiene a  $P_k$

$f_i$  es la frecuencia absoluta del intervalo que contiene a  $P_k$ .

$F_{i-1}$  es la frecuencia acumulada absoluta del intervalo inmediatamente anterior al intervalo que contiene a  $P_k$ .

$A$  es la amplitud del intervalo.

### EJEMPLO

Para los 45 ingresos quincenales tabulados en la distribución de frecuencias de 8 intervalos los datos recopilados fueron:

63	89	36	49	56	64	59	35	78
43	53	70	57	62	43	68	62	26
64	72	52	51	62	60	71	61	55
59	60	67	57	67	61	67	51	81
53	64	76	44	73	56	62	63	60

Se tienen los percentiles:

$$P_{10} = L_i + \left[ \frac{n(10/100) - F_{i-1}}{f_i} \right] A = 42 + \left[ \frac{4.5 - 3}{4} \right] 8 = 42 + 3 = 45$$

$$P_{25} = Q_1 = L_i + \left[ \frac{n(25/100) - F_{i-1}}{f_i} \right] A = 50 + \left[ \frac{11.25 - 7}{10} \right] 8 = 50 + 3.4 = 53.4$$

$$P_{50} = Q_2 = Me = L_i + \left[ \frac{n(50/100) - F_{i-1}}{f_i} \right] A = 58 + \left[ \frac{22.5 - 15}{16} \right] 8 = 60.75$$

$$P_{75} = Q_3 = L_i + \left[ \frac{n(75/100) - F_{i-1}}{f_i} \right] A = 66 + \left[ \frac{33.25 - 33}{10} \right] 8 = 66.75$$

$$P_{90} = L_i + \left[ \frac{n(90/100) - F_{i-1}}{f_i} \right] A = 66 + \left[ \frac{4.5 - 3}{8} \right] 8 = 66 + 7.5 = 72.5$$

## ACTIVIDADES

1. En una prueba de aptitud mental la menor y mayor puntuación fueron 50 y 199 respectivamente. Los puntajes (sin decimales) se tabularon en una distribución de frecuencias simétrica de 5 intervalos de igual amplitud donde el 20% de los casos son menores de 95 y el 70% de los casos son menores que 140. Hallar el intervalo centrado en la mediana donde se encuentran el 50% de los puntajes.}

2. El consumo mensual de agua (en metros cúbicos) de una muestra de 225 viviendas, se tabularon en una distribución de frecuencias simétrica de cinco de amplitud iguales. Si el consumo mínimo es de  $35 \text{ m}^3$ , el consumo promedio de  $45 \text{ m}^3$ , y si  $1/3$  de la muestra consume al menos  $43 \text{ m}^3$  pero menos de  $47 \text{ m}^3$ .

- a) ¿Qué porcentaje de la muestra consume al menos  $47 \text{ m}^3$ ?  
b) ¿Cuántos metros cúbicos como mínimo consumen el 60% de las viviendas?

3. Los porcentajes de artículos defectuosos encontrados en un número determinado de cajas recibidas varían de 10 a 25 y han sido tabulados en una distribución de frecuencias relativas respectivas del primero al tercero 0,08, 0,24; 0,36. Una caja se considera óptima si no supera el 20%.
- a) Halle el porcentaje de cajas óptimas y casi óptimas.  
b) Si las utilidades por cada caja es de 30 unidades monetarias (u.m.) para las óptimas, 15 u.m. para las casi óptimas y 5 u.m. para el resto, ¿Cuánto es la utilidad promedio por caja?



4. Tres obreros utilizaron 480, 360, 240 minutos respectivamente para hacer cierto tipo de objetos. Si utilizaron 0,8; 1; 1,5 minutos por objeto, calcular el tiempo promedio por objeto.

## MEDIDAS DE VARIABILIDAD O DISPERSION

**Medidas de Variabilidad o Dispersión:** Son las que se utilizan para analizar el grado de heterogeneidad de un conjunto de datos. El grado de variabilidad de la información disponible es muy importante en todo análisis estadístico pues de esto depende el grado de confiabilidad de las estimaciones que se puedan establecer. Las medidas de variabilidad que estudiaremos son la varianza o variancia, la desviación estándar y el coeficiente de variación.

Para comparar la dispersión de dos conjuntos de datos es preferible utilizar el coeficiente de variación.

Cálculo de las medidas de variabilidad o dispersión:

a) **Varianza o Variancia (V):**

Para datos no agrupados	Para datos agrupados	
	Discretos	Continuos
$V = \frac{\sum (x_i)^2 - n(\bar{x})^2}{n^*}$	$V = \frac{\sum (x_i)^2 \cdot f_i - n(\bar{x})^2}{n^*}$	$V = \frac{\sum (x_i^1)^2 \cdot f_i - n(\bar{x})^2}{n^*}$

Donde:  $n^* = n - 1$  si se trata de una muestra (n: tamaño de la muestra)

$n^* = n$  si se trata de una muestra (n: tamaño de la población)

b) **Desviación estándar (S)**: Es la raíz cuadrada de la varianza

c) **Coeficiente de Variación (C.V.)**: Se expresa en porcentaje y es el cociente de la desviación estándar y la media aritmética.

$$C.V. = \left( \frac{S}{\bar{x}} \right) (100\%)$$

## ACTIVIDADES

1. Si en una empresa B de 100 trabajadores el sueldo promedio es 500 soles, con una varianza de 900 soles, y en la empresa C el coeficiente de variación de los sueldos es del 5.6% ¿Qué podría afirmar acerca de la dispersión de los sueldos de las empresas C y B?

2. Los gastos mensuales en el uso de Internet de dos empresas se muestran a continuación:  
 Empresa A: Gasto promedio: 1 350 soles, Desviación estándar: s/.78.  
 Empresa B: Gasto promedio: 1 450 soles, Desviación estándar: s/.120.  
 Determine qué empresa tiene gastos mensuales más homogéneos. Justifique.

3. Una empresa de estudios de mercado ha realizado de un trabajo, para estudiar, entre otras variables el pago mensual por alquiler de departamentos (con características similares) en los distritos A y B. La información recogida fue la siguiente:

Pago mensual por alquiler (en \$)	Número de Departamentos (Dist. A)	Número de Departamentos (Dist. B)
[350, 400[	10	9
[400, 450[	35	36
[450, 500[	70	60
[500, 550[	40	75

[550, 600[	25	90
[600, 650[	15	24
[650, 700]	5	6

Con esta información se calculó los siguientes valores para el Distrito A

Medida Estadística	Valor
Media Aritmética	500
Moda	476,92
Variancia	4773,87
Primer Cuartil	457,14
Segundo Cuartil	489,29

- Identifique la población, muestra, variables y tipos de variable que puede encontrar en el estudio.
- Realice un análisis comparativo de las distribuciones de montos mensuales por alquileres de departamentos de los dos distritos.
- Según los contratos de alquiler, el impuesto municipal (15%) debe ser pagado por el dueño del departamento, del mismo modo éste también paga el servicio de mantenimiento del edificio \$30. ¿Cuál es el valor del promedio aritmético, varianza, coeficiente de variabilidad de los ingresos netos por alquiler que tienen los dueños de los departamentos? Explique.
- Se considera que un distrito es más residencial aquel cuyo pago mensual por alquiler es más del 40% de los datos observados. ¿Cuál de los dos distritos se podría considerar más residencial?

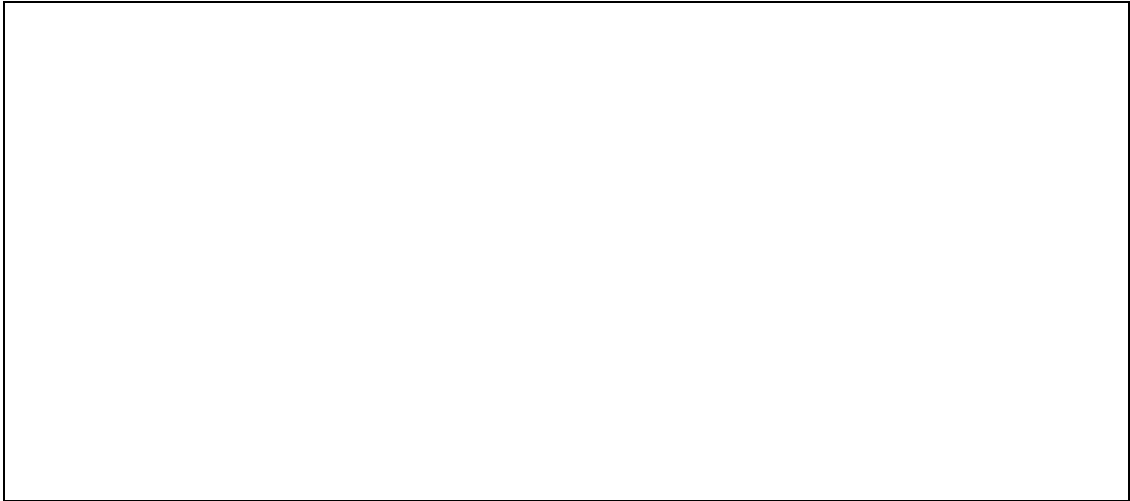
- Un conjunto habitacional está conformado por tres edificios de departamentos. Se tiene los siguientes datos respecto al consumo mensual de electricidad de cada uno de los edificios:  
 Edificio 1: Tiene 12 departamentos que gastan en promedio 45 soles con una desviación estándar de 10 soles.  
 Edificio 2: Tiene 9 departamentos cuyos consumos en soles son: 38, 42, 56, 60, 43, 52, 41, 44, 53.  
 Edificio 3: Los consumos se dan en la siguiente tabla:

Consumo en soles	Departamentos
[30, 40[	1
[40, 50[	4
[50, 60[	4
[60,70]	2

- a) ¿Cuál de los edificios tiene menor consumo promedio de electricidad?
- b) ¿Cuál es el consumo promedio en todo el conjunto habitacional?
- c) ¿En cuál de los edificios el consumo es más disperso?

5. El 20% de los alumnos de una sección tienen 17 años, el 50% tienen como máximo 18 años, el 75% tienen como máximo 19 años y el resto tienen 20 años. Halle la edad promedio de la sección.

6. El 40% de los sueldos de los empleados de una empresa es mayor o igual a 50 soles pero menor de 60 soles; el 30% mayor o iguales a 60 soles pero menor de 70 soles; el 15% de los empleados tienen como mínimo sueldos de 70 soles pero menores de 80 soles y los sueldos del 15% restante son mayores o iguales a 80 soles pero como máximo 100 soles. Halle la media aritmética de los sueldos de los empleados.



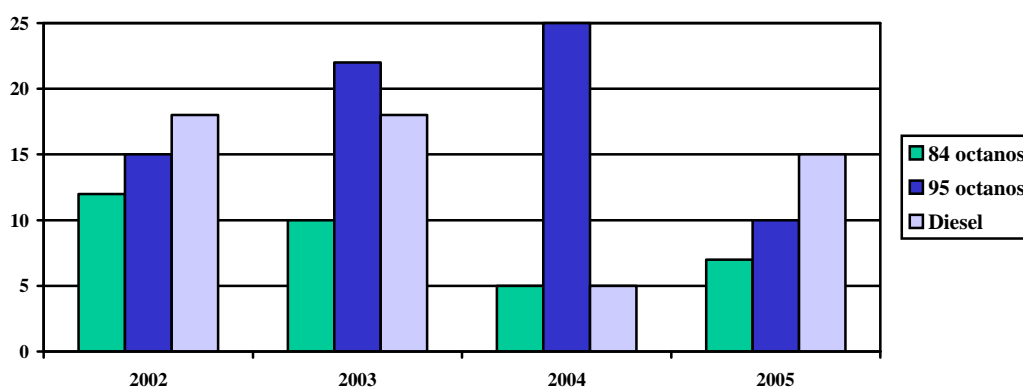
## Autoevaluación

1. Los pobladores de 6 pueblos: A, B, C, D, E, F ubicados a lo largo de la carretera marginal de la selva y en línea recta, desean construir una escuela. Si la población escolar es: 5% de A, 20% de B a 15 Km de A, 30% de C a 20 Km de B, 20% de D a 15 Km de C, 10% de E a 8 Km de D y 15% de F a 6 Km de E, ¿En qué lugar debe construirse la escuela? de manera que:
  - 1.1 El total de las distancias sea mínima
  - 1.2 El costo total de transporte sea mínimo si el costo del transporte es proporcional al cuadrado de la distancia
2. Los salarios que ofrece una empresa a los practicantes varían entre \$150 y 270\$. Si los salarios se agrupan en cuatro intervalos de clase de longitudes iguales de manera que el 40% des los practicantes tienen salarios menores o iguales que \$195, el 80% tienen salarios menores o iguales que \$225 y el 15% tienen salarios mayores que \$232,50.
  - 2.1 ¿Qué porcentaje de practicantes tiene un salario superior al salario medio?
  - 2.2 Si el ingreso mínimo se fija en \$240 y la empresa documenta una misma cantidad a todos los practicantes de modo que el 20% supere el ingreso mínimo, ¿cuánto sería el aumento?
3. Si la población de una ciudad fue de 10 000 habitantes en 1960, de 40 000 en 1980 y de 640 000 en 2000. Calcule la tasa de cambio promedio.
4. Un ahorro de \$1000 acumula intereses variables de 3%, 5%, 8%, durante 3 años, Calcule la tasa de crecimiento promedio.
5. Una prueba de conocimientos. A. se calificó sobre 20 puntos dando una media de 12 y una desviación estándar de 2 puntos. Mientras que una prueba de aptitud. B, se calificó sobre 100 puntos, con una media de 70 y una desviación estándar de 5.
  - 5.1 ¿En cuál de las dos pruebas los puntajes son más homogéneos?
  - 5.2 Si Juan tiene 14 en A y Luís 73 en B. ¿quién tiene mejor rendimiento?
6. Los sueldos de 150 trabajadores de una empresa tienen un coeficiente de variación del 5% en el mes de agosto. Para el mes de septiembre hay un aumento a cada trabajador del 20% de su sueldo más una bonificación de \$60 y el coeficiente de variación baja a 4%.
  - 6.1 Halle la media y la desviación estándar de los sueldos del mes de agosto.
  - 6.2 ¿Cuánto dinero adicional necesita la empresa para pagar todos los sueldos del mes de septiembre?
7. El costo de producción  $X$  de una muestra de cierto tipo de objeto tiene una desviación estándar de \$30. El costo medio de producción es de \$250 para el 60% de la muestra y de \$200 para el resto. Si su precio de venta en dólares es dado por la relación  $Y = 1.1X + 10$ . Calcule la media y la varianza de la venta de la muestra.
8. - La distribución de los sueldos (en dólares) de los empleados de dos empresas A y B se tabuló en 3 intervalos de igual amplitud en cada caso, siendo las frecuencias absolutas del primero al tercero de 10, 30, 30 y de 30, 50, 20 respectivamente en A y B. Si los sueldos mínimo y máximo son de 50 y 200 en A y de 60 y 240 en B.
  - 8.1 ¿En qué empresa los sueldos son más homogéneos?
  - 8.2 Si un empleado de A y otro de B ganan cada uno \$130. ¿Quién de ellos está mejor considerado en su centro de trabajo?

9. El quinto ciclo de una Facultad de Ingeniería tiene tres secciones. La sección A tiene 45 alumnos y la nota promedio en el curso de matemática fue de 13. La sección B tiene 40 alumnos y la nota promedio en el mismo curso fue de 12. Por último la sección C tiene 35 alumnos y la media aritmética de las notas en el curso de matemática fue de 15. Hallar el promedio final de del quinto ciclo de la mencionada Facultad de ingeniería.
10. Una institución educativa tiene 20 profesores. La edad promedio de los profesores casados, que son 12, es de 36 años. La edad promedio de todos los profesores de la institución es 30 años. Calcule la edad promedio de los profesores solteros.
11. En una empresa el sueldo promedio de 60 trabajadores administrativos es 1200 soles. Por incremento del costo de vida se presentan dos alternativas de aumento. La primera propuesta es un aumento de 180 soles a cada trabajador y la segunda es un aumento de 10% de sus sueldos más 12 soles. ¿Cuál de las dos propuestas conviene más a los trabajadores a fin de mejorar su ingreso promedio? Justifique su respuesta.
12. El consumo mensual de leche en litros en los hogares del pueblo ZZZ se muestran a continuación:

Consumo de leche	Cantidad de familias
[0; 8[	6
[8; 16[	9
[16; 24[	23
[24; 32[	40
[32; 40]	42

- 12.1 ¿Cuál es el consumo mensual representativo de las familias del pueblo ZZZ? Justifique su respuesta.
- 12.2 Si el consumo promedio de las 100 familias del pueblo YYY es de 22 litros mensuales, ¿Cuál será el consumo promedio de los dos pueblos?
13. Se presenta la siguiente información acerca de la venta de combustible (en cientos de miles de galones) entre los años 2002 y 2005 en la estación de servicio "PQR"



Determine:

- 13.1 ¿Cuál es el combustible que más demanda tuvo en el periodo de 2002 y 2005? ¿Por qué?
- 13.2 ¿En qué año la demanda fue mayor? ¿Por qué?
- 13.3 ¿Entre qué años se produjo un mayor incremento o disminución en la demanda de combustible? ¿Por qué?



## Para recordar

- Los percentiles nos brindan la misma información que nos proporcionaba la mediana.
- Las medidas de posición son útiles para establecer valores mínimos o máximos que se deben lograr como pautas para mejorar el rendimiento de la producción, de las personas, etc.
- Una mayor dispersión de datos implica una menor concentración de los mismos y viceversa.
- Una mayor heterogeneidad (dispersión) en los datos equivale una menor homogeneidad de los mismos y viceversa.
- A mayor coeficiente de variación, mayor dispersión y, por lo tanto, menos homogeneidad de los datos.
- El CV. nos permite determinar la representatividad del promedio de un conjunto de datos, pues si es menor del 50% podemos considerar al promedio como representativo de los datos. En caso contrario se considera que los datos son muy dispersos y, por lo tanto, no es recomendable utilizarla.





# ANÁLISIS COMBINATORIO PROBABILIDADES

## TEMAS

- Análisis combinatorio
- Probabilidad. Definiciones generales

## OBJETIVOS ESPECÍFICOS

- Utilizar en forma adecuada las combinaciones y permutaciones para resolver problemas de conteo.
- Comprenderá el concepto de probabilidad y demás definiciones relacionadas con dicho concepto.

## CONTENIDOS

- Análisis combinatorio
- Principio de la adición.
- Principio de la multiplicación.
- Permutaciones
- Combinaciones
- Experimento aleatorio.
- Espacio muestral
- Suceso o evento

## ACTIVIDADES

- Calculan el número de eventos que tiene un espacio muestral
- Manejan las definiciones básicas de probabilidad

## Análisis Combinatorio

Es el conjunto de procedimientos que permiten determinar el número de resultados de un experimento sin necesidad de conocer todos los resultados que de él se originan. Se basa en los siguientes principios:

1. **Principio de la adición:** Dados dos experimentos A y B, el número de maneras posibles que puede ocurrir el experimento A o B está dado por:

$$n(A \text{ o } B) = n(A) + n(B)$$

Donde:

$n(A)$ : Número de formas distintas que puede ocurrir el experimento A.

$n(B)$ : Número de formas distintas que puede ocurrir el experimento B.

2. **Principio de la multiplicación:** Dados dos experimentos A y B, el número de maneras posibles que puede ocurrir el experimento A y B está dado por:

$$n(A \text{ y } B) = n(A) \cdot n(B)$$

Ejemplo: Para adquirir un equipo de cómputo que conste de teclado, monitor, CPU e impresora se cuenta con 4 marcas distintas de teclado, 5 marcas distintas de monitor, 3 de CPU y 4 de impresoras. Suponiendo que todas las marcas son compatibles determine de cuántas maneras es posible adquirir el equipo de cómputo.

3. **Permutaciones:** La permutación de un conjunto de elementos es un arreglo de los mismos siguiendo un orden establecido. Se pueden presentar los siguientes casos:

- 3.1 El número de permutaciones posibles de “n” elementos, todos distintos, agrupados en sub grupos de “n” elementos diferentes son:

$$P_n^n = n!$$

- 3.2 El número de permutaciones posibles de “n” elementos, todos distintos, agrupados en sub grupos de “r” elementos diferentes son:

$$P_r^n = \frac{n!}{(n-r)!}$$

4. El número de permutaciones posibles de “n” elementos, todos distintos, agrupados en sub grupos de “r” elementos no necesariamente distintos (con repetición)son:

$$P_{r(\text{repetición})}^n = n^r$$

- 5.: El número de permutaciones posibles de “n” elementos, no todos distintos y donde se distingue  $n_1$  elementos iguales,  $n_2$  elementos iguales, hasta  $n_k$  elementos iguales ( $n_1 + n_2 + n_3 + \dots + n_k = n$ ); agrupados en sub grupos de “n” elementos es:

$$P_{n_1, n_2, \dots, n_k}^n = \frac{n!}{n_1! n_2! n_3! \dots n_k!}$$

5. **Combinaciones:** Una combinación de un conjunto de elementos es una selección de tales elementos sin tener en cuenta el orden. El número de combinaciones de “r” elementos tomados de “n” en “r” (sub grupos de “r” elementos) es:

$$C_r^n = \frac{n!}{r!(n-r)!}$$

## ACTIVIDADES

1. ¿De cuántas maneras se pueden colocar 10 chicas en una fila, de manera que dos chicas en particular queden juntas?

2. ¿De cuántas formas pueden 5 ordenarse personas en un sofá si tienen solamente tres asientos?

3. ¿De cuántas formas puede seleccionarse 6 preguntas de un total de 10?

4. ¿Cuántos comités diferentes de 3 hombres y 4 mujeres pueden formarse con 8 hombres y 6 mujeres?

5. ¿De cuántas maneras se puede confeccionar una bandera de franjas de 3 colores, si se tiene tela de 5 colores distintos? , ¿si una de las franjas tiene que ser azul?

6. Una persona tiene 15 amigos. De cuántas maneras puede invitar a 6 de ellos.  
6.1 Si existen 2 parejas de casados y van siempre juntos a cualquier reunión.  
6.2 Si existen dos personas que no deben estar en la misma reunión.

- 7 En una clínica trabajan 18 enfermeras.  
7.1 Cuantas guardias diferentes de 3 enfermeras pueden formarse.  
7.2 Cuantas guardias diferentes de 3 enfermeras pueden formarse, si dos ellas no pueden trabajar juntas

8. Se ha comprado una caja de focos que contiene, 2 de 25 vatios, 3 de 50 vatios, y 4 de 100 vatios. Si se escogen 3 de ellos, ¿cuántos incluirán uno de cada potencia?

9. Diez amigos se encuentran en una fiesta, ¿Cuántos saludos de mano se intercambian, si cada amigo estrecha la mano a todos los demás sólo una vez?

10. En una biblioteca hay 8 libros de geometría, 14 de álgebra, 10 de física, y 5 de química. ¿De cuántas maneras un estudiante debe seleccionar cuatro libros, si estos deben ser de diferentes materias?



## Definiciones Generales

**Experimento Aleatorio:** Es cualquier experimento cuyo resultado inmediato no se puede predecir con certeza y que puede ser repetido bajo las mismas condiciones. Sin embargo es posible conocer, todos los posibles resultados que se pueden presentar.

**Espacio Muestral ( $\Omega$ ):** Es el conjunto formado por todos los posibles resultados de un experimento aleatorio. Puede ser: Finito o Infinito, y este último, numerable o innumerable

**Suceso o Evento:** Es cualquier sub conjunto del espacio muestral.

**Sucesos mutuamente excluyentes:** Son aquellos que no pueden darse simultáneamente. La ocurrencia de uno imposibilita la ocurrencia del otro. Si A y B son mutuamente excluyentes se cumple que:  $A \cap B = \emptyset$

## Actividades

1. Describe el espacio muestral para cada uno de los eventos siguientes:
  - 1.1 Tres lanzamientos de una moneda.
  - 1.2 El número de fumadores de un grupo de 500.

2. Describe el espacio muestral para cada uno de los eventos siguientes:
  - 2.1 El lanzar una moneda hasta que aparezca un sello.
  - 2.2 El lanzar una moneda y un dado

3. Describe el espacio muestral para cada uno de los eventos siguientes:
  - 3.1 Lanzar dos dados y observar su resultado.
  - 3.2 Promedio de vida de un artefacto eléctrico.

4. Describe el espacio muestral para cada uno de los eventos siguientes:
- 4.1 La posición de caída al disparar a un blanco circular de 25 centímetros de radio.
  - 4.2 El agrupar tres lápices de color de un total de 6 diferentes.

5. Una urna contiene dos bolas B1 y B2 y dos bolas negras N1 y N2. Se realiza el siguiente experimento: se sacan dos bolas en forma sucesiva y sin restitución.

- 5.1 Define el espacio muestral asociado al experimento descrito
- 5.2 Describe el siguiente evento: "las dos bolas extraídas son blancas"

# Autoevaluación

1. En un ómnibus que posee 37 asientos (en 8 filas de cuatro asientos cada una con un pasillo en el medio y al final 5 asientos juntos), se desea ubicar 25 pasajeros.
  - 1.1 ¿De cuántas formas se pueden ubicar?
  - 1.2 ¿De cuántas formas se pueden ubicar si deciden no ocupar los 5 últimos asientos?
  - 1.3 ¿De cuántas formas se pueden ubicar si viajan 5 amigos que deciden viajar juntos en los últimos asientos?
  - 1.4 ¿De cuántas formas se pueden ubicar si ocupan los 18 asientos que poseen ventanilla?
  - 1.5 ¿De cuántas formas se pueden ubicar si 10 de los pasajeros están enfermos y deben viajar en asientos que poseen ventanilla?
2. De cuántas formas pueden ordenarse 7 libros en un estante sí:
  - 2.1 es posible cualquier ordenación
  - 2.2 Tres libros determinados deben estar juntos
  - 2.3 Dos libros determinados deben ocupar los extremos.
3. Cuántos números de 5 dígitos puede formarse con los 10 dígitos del sistema decimal:
  - 3.1 si se permiten repeticiones
  - 3.2 sin repeticiones
4. Un entrenador de volley dispone de 6 jugadoras veteranas y de 5 novatas para hacer la alineación de su equipo. Si las jugadoras pueden jugar en cualquier posición, ¿cuántas alineaciones se podrán hacer, si en el equipo necesariamente debe haber 3 jugadoras veteranas y 3 novatas?
5. El asta de bandera de un barco tiene tres posiciones en las que se pueden colocar una bandera en cada una. Si el barco lleva cuatro banderas diferentes para hacer señales. Cuantas señales diferentes se pueden hacer:
  - 5.1 Con una bandera.
  - 5.2 Con dos banderas.
  - 5.3 Con las banderas.
6. De un grupo de 10 son hombres y 15 mujeres; se quiere formar grupos de 5 personas
  - 6.1 Que exista dos hombres en cada grupo.
  - 6.2 Que los grupos estén formados sólo de igual sexo.
7. Un niño tiene 8 bolas blancas y 2 bolas negras, las alinea al azar,
  - 7.1 ¿De cuántas formas las de color negro queden juntas?
  - 7.2 ¿De cuántas formas las de color negro ocupen posiciones de los extremos?

- 8 ¿Cuántas cantidades diferentes de dinero pueden formarse con; una moneda de 50 céntimos, una moneda de 1 sol, una moneda de 5 soles, un billete de 10 soles, un billete de 50 soles, y un billete de 100 soles?
- 9 ¿Cuántos equipos de fútbol pueden formarse con 12 hombres que pueden ocupar cualquier posición delantera, y 10 hombres que pueden ocupar cualquiera de las demás posiciones?

## Para recordar

- En el análisis combinatorio una conjunción (y) implica una multiplicación; en cambio, una disyunción (o) implica una suma.
- Debemos utilizar permutaciones si queremos permutar, ordenar, arreglar, colocar, situar, ubicar, etc.; un conjunto de datos.
- Debemos utilizar combinaciones si debemos combinar, escoger, seleccionar, elegir, etc.
- La mayoría de las actividades que realiza el hombre se puede considerar como un experimento aleatorio: la duración de las llantas de una determinada marca, la respuesta que va dar una persona encuestada, el número de hijos que pueda tener una familia cuya vivienda pertenece a una calle determinada, etc.
- Para establecer el número de elementos del espacio muestral se utiliza el análisis combinatorio.
- Por lo general no es necesario saber “cuáles” son los elementos del espacio muestral, sino “cuántos” elementos son.







SEMANA

6

# PROBABILIDADES

## PROBABILIDAD CONDICIONAL

### TEMAS

- Álgebra de eventos. Propiedades
- Probabilidad condicional. Regla del producto

### OBJETIVOS ESPECÍFICOS

- Relaciona los distintos eventos mediante conectivos adecuados y calcular la probabilidad asociada a ellos.
- Utiliza en forma adecuada los criterios inherentes a la probabilidad condicional y a la probabilidad en condiciones de dependencia estadística.

### CONTENIDOS

- Definición clásica de probabilidad
- Definición axiomática de probabilidad.
- Propiedad y teoremas de probabilidad.
- Probabilidad Condicional
- Teorema de la multiplicación de probabilidad

### ACTIVIDADES

- Resuelven problemas de probabilidad
- Resuelven problemas de probabilidad condicional

# Probabilidad

Es un número real que expresa la confianza o incertidumbre en la ocurrencia de un evento cuyo resultado no se puede predecir con certeza.

**Definición Clásica de Probabilidad:** Si un experimento aleatorio se puede realizar de “n” maneras posibles y mutuamente excluyentes y  $n_A$  de ellos tiene una característica A, la probabilidad que se obtenga un resultado con característica A es:

$$P(A) = \frac{n_A}{n}$$

**Definición axiomática de probabilidad:** Sea un experimento aleatorio E con espacio muestral  $\Omega$  y A un evento cualquiera de  $\Omega$ . El número real  $P(A)$  es llamado probabilidad de ocurrencia del evento A, si satisface las siguientes condiciones:

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

$$P(\Omega) = 1$$

Si A y B son mutuamente excluyentes, se cumple que:

$$P(A \cup B) = P(A) + P(B)$$

## Propiedades y teoremas básicos de probabilidades

Dados tres eventos A, B y C contenidos en el espacio muestral  $\Omega$  se cumple:

- a)  $P(\phi) = 0$
- b)  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- c)  $P(A') = 1 - P(A)$
- d)  $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

## Actividades

1. Con 7 abogados y 4 médicos se van a formar comités de 6 personas. Hallar la probabilidad de que se forme un comité en el cual se incluya:
  - a) Exactamente 2 médicos
  - b) Por lo menos 2 abogados
  - c) Como máximo 3 médicos

2. Una empresa que ofrece servicios de soporte informático cuenta con 5 profesionales que sólo manejan Visual Basic, 4 manejan sólo Unix y 3 que

manejan Windows NT. Una compañía le solicitó que envíen un equipo de tres personas. Calcule la probabilidad que el equipo esté formado por:

- a) Una persona que maneje sólo VB, otra Unix y la otra Windows NT.
- b) Personas que sólo manejan una sola especialidad.
- c) 2 personas que sólo manejan Windows NT.

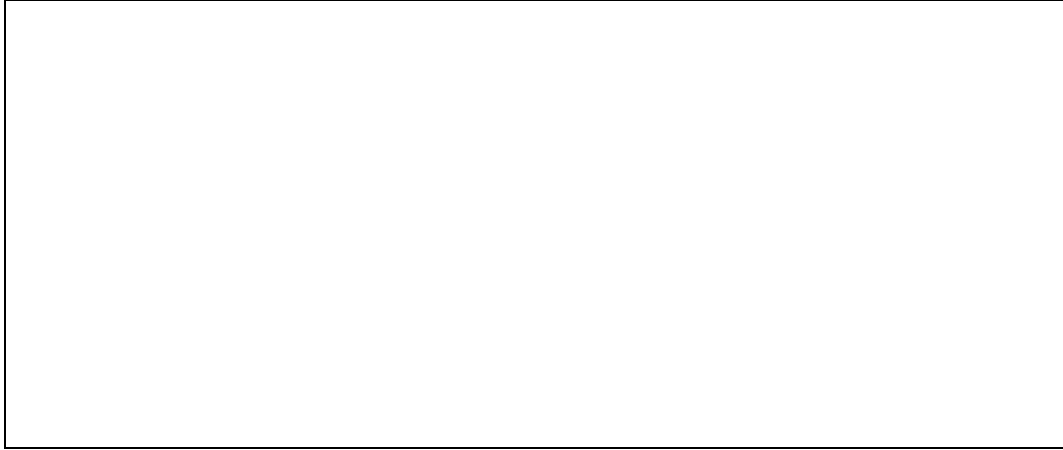
3. De 20 personas que contrajeron cierta enfermedad al mismo tiempo y que fueron llevados a una misma sala de hospital, 15 se recuperan completamente en 3 días; al cabo del cual, se escogen aleatoriamente 5 personas para un chequeo.

- a) ¿Cuál es la probabilidad que los cinco sean dados de alta?,
- b) ¿Cuál es la probabilidad que exactamente cuatro sean dados de alta?,
- c) ¿Cuál es la probabilidad que ninguno sea dado de alta?

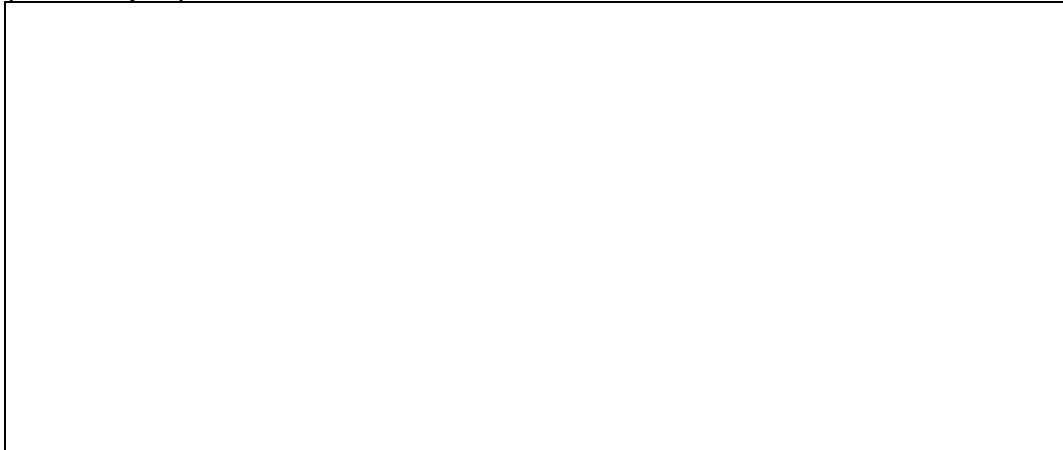
4. Ocho parejas de casados se encuentran en un salón. Si se escogen 2 personas al azar, hallar la probabilidad de que sean:

- a) Esposos.
- b) Una sea mujer y el otro hombre.

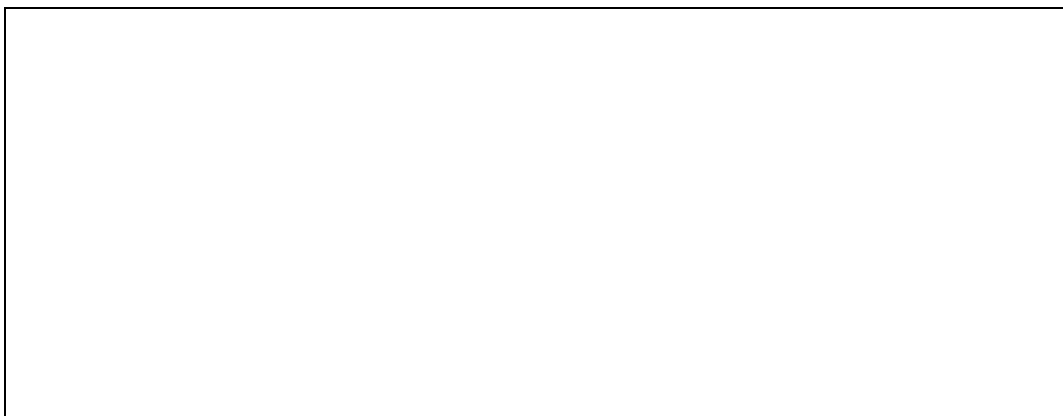
5. Una urna A contiene 3 bolas rojas y 3 bolas negras, mientras que la urna B contiene 4 bolas rojas y 6 negras. Si una bola es extraída aleatoriamente de cada urna. ¿Cuál es la probabilidad de que las bolas sean del mismo color?



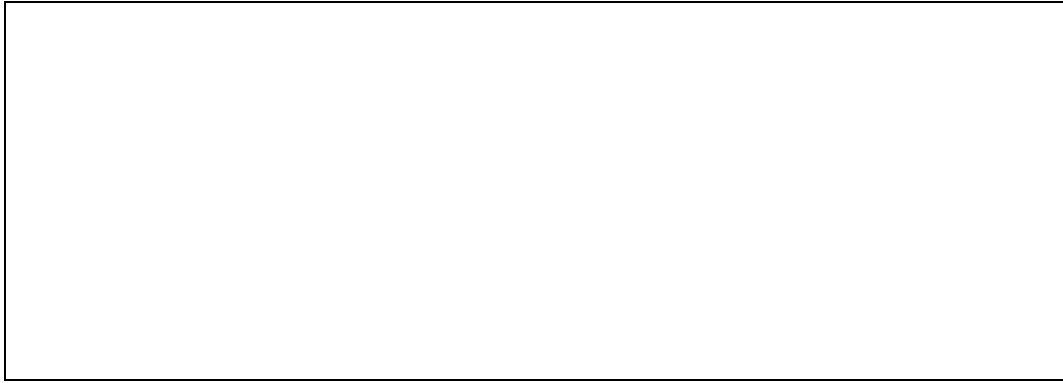
6. En una caja hay 7 bolas rojas y 3 verdes, se sacan dos bolas al azar de una en una. Hallar la probabilidad de que la primera sea roja y la segunda también si es que no hay reposición.



7. Ocho ejecutivos de una empresa llegan diariamente a su oficina en un automóvil y lo aparkan en una de las tres playas de estacionamiento con que cuentan. Si los estacionamientos son escogidos al azar, Cuál es la probabilidad de que en un día determinado se tenga 5 automóviles en un estacionamiento, dos en otro y el restante en el otro.



8. En una oficina donde se ensamblan computadoras, en una mesa hay 20 chips de los cuales 6 están malogrados. Primero llega el Sr. Gates y recoge 8 chips y más tarde llega el Sr. Apple y se lleva los restantes. Halle la probabilidad que solamente uno de ellos se haya llevado todos los chips defectuosos.



## PROBABILIDAD CONDICIONAL

1. **Probabilidad Condicional:** Si A y B son dos eventos de un espacio muestral  $\Omega$ , entonces la probabilidad condicional que ocurra el evento A dado que ocurrió el evento B es:

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

2. **Teorema de la multiplicación de probabilidades**

Si dados tres eventos A, B y C contenidos en el espacio muestral  $\Omega$  se cumple:

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

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

## Actividades

1. Cierta universidad en formación en su primer año de funcionamiento tiene tres currículos: Ciencia, Administración e Ingeniería. La clasificación de alumnos por sexo es así:

	<b>Ciencia</b>	<b>Administración</b>	<b>Ingeniería</b>	<b>Total</b>
Hombres	250	350	200	800
Mujeres	100	50	50	200
Total	350	400	250	1000

Se selecciona un estudiante aleatoriamente del grupo. Si se sabe que el estudiante puede ser hombre o mujer:

- 1.1 ¿Cuál es la probabilidad que esté en Ciencias, si es varón?
- 1.2 Si es mujer, ¿Cuál de que esté en Ingeniería.
- 1.3 ¿Cuál es la probabilidad que el estudiante esté matriculado en Administración?, Dado que es mujer.

2. En una universidad el 70% de los estudiantes son de Ciencias y el 30% de Letras; de los estudiantes de Ciencias el 60% son varones y los de letras son varones el 40%. Si se elige aleatoriamente un estudiante, calcular la probabilidad que sea un estudiante:

- 2.1 Varón
- 2.2 Varón, si es de Ciencias
- 2.3 Ciencias, si es varón
- 2.4 Ciencias y varón

3. Suponga en una oficina que hay 100 calculadoras de las cuales 60 son eléctricas y las otras son manuales. Además hay 70 máquinas nuevas

mientras que las otras son usadas. Hay 30 máquinas nuevas y manuales. Si se escoge una máquina al azar:

3.1 ¿Cuál es la probabilidad de que sea eléctrica?

3.2 ¿Cuál es la probabilidad que sea manual si se sabe que es usada?

4. La probabilidad de que un vuelo de programación regular despegue a tiempo es 0,83, que llegue a tiempo es 0,82 y de que despegue y llegue a tiempo es 0,78. Encuentre la probabilidad que un avión:

4.1 Llegue a tiempo dado que no despegó a tiempo

4.2 despegue a tiempo dado que llegó a tiempo

5. Un hombre tiene dos carros A y B que tienen problemas de arranque. La probabilidad de que ambos arranquen es 0,1, la probabilidad de que arranque B y no A es 0,2, la probabilidad de que ninguno de ellos arranque es 0,4. Halle la probabilidad de que:

5.1 El carro A arranque

5.2 Arranque A dado que B no arrancó

6. La probabilidad de que un alumno apruebe Estadística es de  $\frac{2}{3}$  y la probabilidad de que aprueba Matemática es  $\frac{4}{9}$ . Si la probabilidad de aprobar al menos una de estas materias es  $\frac{4}{5}$ . ¿Cuál es la probabilidad de que apruebe ambos cursos?

7. Sean A y B dos eventos tales que:  
 $P(A) = 0.2$ ,  $P(B') = 0.4$ ,  $P(A' \cap B') = 0.3$

Calcule:

7.1  $P(A \cup B)$

7.2  $P(A \cap B)$

7.3  $P(A \cap B')$

7.4  $P(A' \cup B)$

8. En el mercado laboral últimamente tienen gran demanda aquellos profesionales con conocimientos avanzados. Se sabe que el 15% de los que solicitan empleo sólo tienen conocimientos de Visual Basic, el 10% sólo conoce Lenguaje C y el 5% sólo conoce Pascal. El 30% no tiene conocimiento de estas tres herramientas de programación. También se sabe que el 35% tiene experiencia solamente en dos de los tres lenguajes de programación.

8.1 Halle la probabilidad de que al entrevistar a un postulante a un puesto de programador, este conozca Visual Basic, Lenguaje C y Pascal.

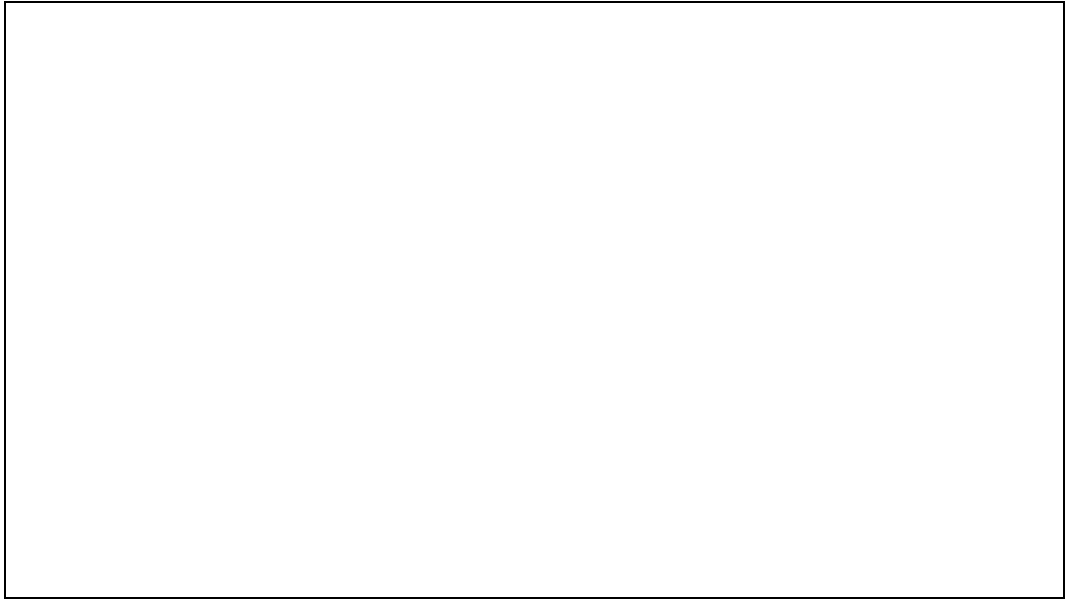


- 8.2 Halle la probabilidad de que un postulante conozca los 3 lenguajes si se sabe que conoce al menos uno.

9. Un alumno de Cibertec está llevando Matemática, Estadística y Contabilidad. Por experiencias anteriores se sabe que la probabilidad de aprobar Matemática es 0,7; el de Estadística 0,55; el de Contabilidad 0,5; el de matemáticas y estadística 0,3; el de matemáticas y contabilidad 0,35; el de estadística y contabilidad 0,3 y los tres a la vez 0,2. Calcule la probabilidad que el alumno:

- 9.1 Apruebe sólo un curso.  
9.2 Apruebe por lo menos dos cursos.  
9.3 No apruebe algún curso

10. En una ciudad el 70% de los adultos escuchan radio, el 40% lee periódico y el 10% ve televisión; entre los que escuchan radio el 30% lee periódicos y el 4% ve televisión. El 90% de los que ven TV lee el periódico y sólo el 2% de la población total lee periódico, ve TV y escucha radio. Si se elige una persona al azar calcule la probabilidad:
- 10.1 Que sólo realice una de las tres actividades.  
10.2 Sabiendo que lee periódico, que vea televisión.  
10.3 Escuche radio o vea TV.



# Autoevaluación

1. Verificar que para eventos A y B cualesquiera:
  - a)  $P(A - B) = P(A) - P(AB)$ .
  - b)  $P(AB^c \cup A^c B) = P(A) + P(B) - 2P(AB)$ .
2. Si A, B y C son sucesos cualesquiera, demostrar que:
 
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$
3. Si la probabilidad de que ocurra un evento A es  $\frac{1}{2}$  y que ocurra un evento B es  $\frac{3}{4}$ , determine los posibles valores de  $p = P(A \cap B)$ .
4. La probabilidad de que Juan asista a una cita es  $\frac{1}{2}$  y de que María asista a la misma cita es  $\frac{5}{14}$ . si la probabilidad de que al menos uno de los dos asista a la cita es  $\frac{5}{7}$ , calcular la probabilidad de que:
  - 4.1 Ninguno de los dos asista a la cita.
  - 4.2 Sólo uno de ellos asista a la cita.
5. Un lote contiene  $n$  objetos. La probabilidad de que al menos uno sea defectuoso es 0.06, mientras que la probabilidad de que al menos dos sean defectuosos es 0.04. calcular la probabilidad de que:
  - 5.1 Todo objeto sean no defectuosos.
  - 5.2 Exactamente un objeto sea defectuoso
6. Una caja contiene 5 fichas de \$10 cada una, 3 de \$30 cada una y 2 de \$50 cada una. Si se escogen 3 fichas al azar y a la vez, calcular la probabilidad de que la suma de los valores sea de \$ 70.
7. Doscientas personas están distribuidas de acuerdo a su sexo y lugar de procedencia de la siguiente manera: 130 son hombres, 110 son de la capital y 30 son mujeres y de provincias. Si se eligen dos personas al azar calcular la probabilidad de que:
  - 7.1 Ambos sean hombres y de provincias.
  - 7.2 Al menos uno de los dos sea mujer.
8. Un lote contiene 8 artículos buenos y 4 defectuosos, si se extraen al azar 3 artículos a la vez, calcular la probabilidad de obtener por lo menos un defectuoso.
9. Una urna contiene 20 fichas similares de las cuales 10 son rojas, 6 son azules, y 4 son verdes. Si se extraen 10 fichas al azar y a la vez, calcular la probabilidad de que.
  - 9.1 Cinco fichas sean rojas.
  - 9.2 Cinco sean rojas, 3 azules; y 2 sean verdes.
10. Un comerciante tiene 12 unidades de cierto artículo de los cuales 4 tiene algún tipo de defecto. Un cliente pide para comprar 3 de tales artículos pero que no

tengan defectos. Si el comerciante escoge al azar y de una sola vez 4 de tales artículos, ¿cuál es la probabilidad de que con las 4 unidades escogidas satisfaga el pedido del cliente?

11. De 6 alumnos de ingeniería y 4 de ciencias se deben seleccionar dos de ellos para hacer cierta tarea, ¿Cuál es la probabilidad de que la selección esté formada por uno de ciencias y otro de ingeniería si un determinado alumno de ciencias no puede hacer pareja con 2 de ingeniería?
12. Una caja contiene 16 pernos de los cuales 8 no tiene defectos, 5 tiene defectos leves, y 3 tiene defectos graves. Si se eligen 3 pernos al azar y de una sola vez, calcular la probabilidad de que los tres no tengan defectos leves.
- 13- Si:  $P(A) = 5/8$ ,  $P(B) = 3/4$  y  $P(A/B) = 2/3$ , calcular  $P(A/B^C)$ .
14. Si:  $P(B) = 3/15$ ,  $P(B/A) = 1/5$  y  $P(A \cap B) = 1/15$ , calcular  $P(B^C/A)$ .
15. En una muestra de 240 Loretanos se encontró que el 60% sufre alguna enfermedad, el 30% tienen al menos 30 años, y el 20% son menores de 30 años y sanos. Si uno de tales loretanos es escogido al azar, cuál es la probabilidad de que:
  - 15.1 Sufra alguna enfermedad y tenga al menos 30 años
  - 15.2 Sufra alguna enfermedad si tiene al menos 30 años
16. De 80 personas encuestadas, se encontró que  $3/4$  son solteras,  $3/8$  son mujeres y uno de cada 10 hombres son casados.
  - 16.1 Si se elige al azar una persona encuestada y resulta ser casada, calcular la probabilidad de que sea un hombre.
  - 16.2 Si se eligen al azar dos personas encuestadas y resultan ser del mismo sexo, calcular la probabilidad de que ambos sean hombres.
17. Del primer examen de estadística resulta que el 40% de los alumnos está en el cuarto superior. Se estima en 70% los alumnos que no estarán en el cuarto superior del segundo examen y una probabilidad de  $3/4$  que un alumno no esté en el cuarto superior del segundo examen si estuvo en el cuarto superior del primer examen. Si se elige al azar un alumno de este curso, calcular la probabilidad de que esté
  - 17.1 En el cuarto superior del segundo examen y no en el cuarto superior del primer examen.
  - 17.2 En el cuarto superior del segundo examen si no estuvo en el cuarto superior del primero.

## Para recordar

- Una probabilidad se puede interpretar como el porcentaje de veces que va ocurrir un determinado evento.
- La probabilidad NUNCA puede ser mayor que uno.
- Cuando la ocurrencia de un evento se puede dar de varias formas, para calcular la probabilidad de ocurrencia de dicho evento es mejor utilizar la probabilidad de ocurrencia del complemento de dicho evento.
- Los eventos tienen un comportamiento similar al de los conjuntos, por ello los diagramas de Venn son bastante útiles en el cálculo de probabilidades.
- En probabilidad condicional el evento que actúa como condicionante es un evento que ya ocurrió.
- Si dos eventos no pueden ocurrir simultáneamente (eventos mutuamente excluyentes) entonces su probabilidad condicional es cero.
- Si queremos calcular la probabilidad de ocurrencia simultanea o sucesiva de una serie de eventos, debemos utilizar el teorema de multiplicación de probabilidades.
- Si la probabilidad de ocurrencia de un evento depende del resultado que se haya dado en un evento anterior, debemos utilizar el teorema de multiplicación de probabilidades.





# PROBABILIDAD TOTAL

## TEOREMA DE BAYES

### TEMAS

- Probabilidad Total Diagrama del Árbol.
- Teorema de Bayes y Independencia Estadística

### OBJETIVOS ESPECÍFICOS

- Comprender en que casos se debe utilizar la regla de probabilidad total y el cuadro de doble entrada, como herramientas en el cálculo de probabilidades.
- Comprender en que casos se debe utilizar la regla de Bayes y diagrama del árbol, como herramientas en el cálculo de probabilidades.

### CONTENIDOS

- Probabilidad Total.
- Partición.
- Regla de Bayes
- Eventos independientes

### ACTIVIDADES

- Resuelve problemas de probabilidad total
- Resuelven problemas de probabilidad condicional aplicando la regla de Bayes

## Probabilidad total

**Partición:** Los eventos  $B_1$ ,  $B_2$  y  $B_3$  forman una partición del espacio muestral si son mutuamente excluyentes y colectivamente exhaustivos, por lo que deben cumplir las siguientes condiciones:

$$B_1 \neq \emptyset; B_2 \neq \emptyset; B_3 \neq \emptyset$$

$$B_1 \cap B_2 = \emptyset; B_1 \cap B_3 = \emptyset; B_2 \cap B_3 = \emptyset$$

$$B_1 \cup B_2 \cup B_3 = \Omega$$

**Probabilidad Total:** Sean los eventos  $B_1$ ,  $B_2$  y  $B_3$  los cuales forman una partición del espacio muestral y sea  $A$  otro evento cualquiera de  $\Omega$ , entonces, la probabilidad de ocurrencia del evento  $A$  está dada por:

$$P(A) = P(B_1).P(A/B_1) + P(B_2).P(A/B_2) + P(B_3).P(A/B_3)$$

### Actividades

1. En un laboratorio hay tres jaulas, la jaula I hay tres conejos negros y dos blancos, en la jaula II tiene dos conejos negros y cuatro blancos y en la jaula tres tiene cinco conejos negros y cinco blancos. Se selecciona al azar una jaula y se saca un conejo al azar de esta jaula. Cuál es la probabilidad que el conejo escogido sea negro.

2. Del record pasado, se conoce que cierta máquina que produce tornillos trabaja correctamente el 90% del tiempo. Si la máquina no trabaja correctamente, el 5% de los tornillos producidos son defectuosos. Si se escoge un tornillo aleatoriamente. Cuál es la probabilidad que este sea defectuoso.



3. Una urna contiene tres bolas rojas y  $X$  blancas. Se extrae una bola de la urna y se reemplaza por una de otro color, se saca de la urna una segunda bola. Sabiendo que la probabilidad de que la segunda bola sea roja es  $17/50$ . Determine el número de bolas blancas.

4. En una urna hay dos bolas rojas y una negra. Hugo, Paco y Luis (en ese orden) deben sacar uno después del otro, una bola sin restituirla posteriormente. ¿Cuál de las tres personas tiene mayor posibilidad de sacar la bola negra?

## Teorema de Bayes

Sean los eventos  $B_1$ ,  $B_2$  y  $B_3$  los cuales forman una partición del espacio muestral y sea  $A$  otro evento cualquiera de  $\Omega$ , la probabilidad de ocurrencia de  $B_k$  ( $k = 1, 2, 3$ ) dado que ocurrió el evento  $A$  está dada por:

$$P(B_k/A) = \frac{P(B_k).P(A/B_k)}{P(B_1).P(A/B_1) + P(B_2).P(A/B_2) + P(B_3).P(A/B_3)}$$

**Eventos independientes:** Dos eventos  $A$  y  $B$  son independientes estadísticamente si la ocurrencia de uno de ellos no influye en la probabilidad de la ocurrencia del otro. Si es así, se debe cumplir:

$$P(A \cap B) = P(A).P(B)$$

## Actividades

1. Una cadena de tiendas de venta de abarrotes desea uniformizar la red interna de cada tienda. Cinco de las tiendas utilizan el sistema Windows NT, 8 de las tiendas usan la red Novell y 7 usan Unix. Se sabe, por experiencia vivida, que en el 5% de las horas de mayor congestión se cae el sistema que está operando con Novell; en el 8% se cae el sistema que está operando con Windows NT y en el 10% se cae el sistema que está operando con Unix.
  - 1.1 Si un día cualquiera, en horas de mayor congestión, se cae el sistema. ¿Cuál es la probabilidad que haya sido en una tienda que está operando con Unix?
  - 1.2 Si ha usted le consultan, acerca de qué sistema deberían utilizar todas las tiendas, a fin de evitar que el sistema se suspenda en las horas de mayor congestión. ¿Qué aconsejaría? Base su respuesta en los datos del problema y en criterios probabilísticos.

2. Una compañía de desarrollo urbano está considerando la posibilidad de construir un centro comercial en un sector de Lima. Un elemento vital en esta consideración es un proyecto de una autopista que une este sector con el centro de la ciudad. Si el Concejo municipal aprueba esta autopista, hay una probabilidad de 0,90 de que la compañía construya el centro comercial en tanto que si la autopista no es aprobada la probabilidad es de solo 0,20. Basándose en la información disponible, se estima que hay una probabilidad de 0,60 que la autopista sea aprobada. Dado que el Centro Comercial fue construido, Cuál es la probabilidad de que la autopista haya sido aprobada.

3. La compañía que ensambla automóviles Neón Génesis, se ha presentado a una licitación para ensamblar un nuevo modelo de automóvil. La probabilidad que Neón

Génesis gane la licitación es de 0,90 si una firma competidora Akira S.A. no se presente a ella, en tanto que es de sólo 0,20 si Akira S.A. se presenta; el gerente general de Neón Génesis estima que hay una probabilidad de 0,80 que Akira S.A. se presente.

3.1 ¿Cuál es la probabilidad que Neón Génesis gane la licitación?

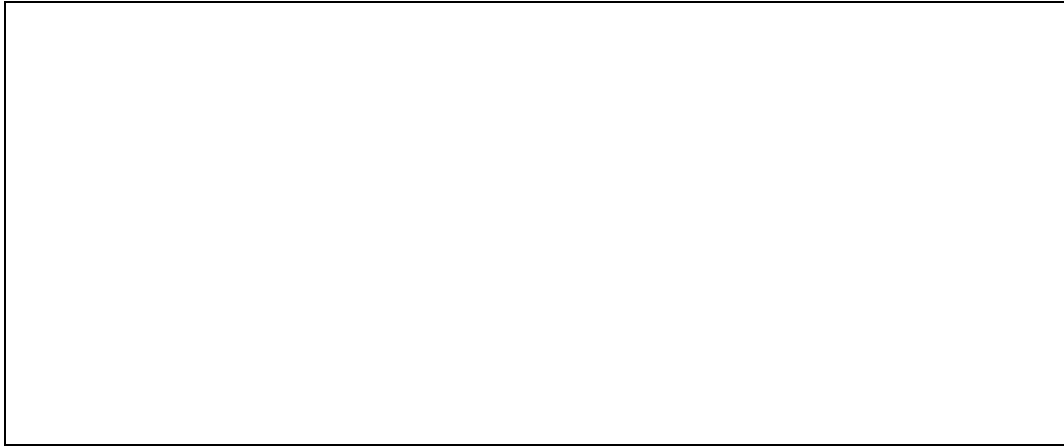
3.2 Dado que Neón Génesis ganó la licitación, ¿cuál es la probabilidad que Akira S.A. se haya presentado a ella?

4. Dos máquinas producen un mismo artículo. La probabilidad de que la máquina 1 produzca un artículo defectuoso es 0,01, mientras que la máquina 2 produce un artículo defectuoso con probabilidad 0,02. De un gran lote de artículos producidos por ambas máquinas se extrae uno al azar. Hallar la probabilidad de que siendo el artículo escogido defectuoso, este provenga de la máquina 1.

5. Se tiene 5 cajas de 100 focos cada una. Dos de las cajas contienen 10 focos defectuosos cada una, las otras dos tienen 5 focos defectuosos cada una y la última caja dos focos defectuosos. Si se selecciona al azar una de estas cajas y de esta se toma un foco:

5.1 ¿Cuál es la probabilidad de que este sea defectuoso?

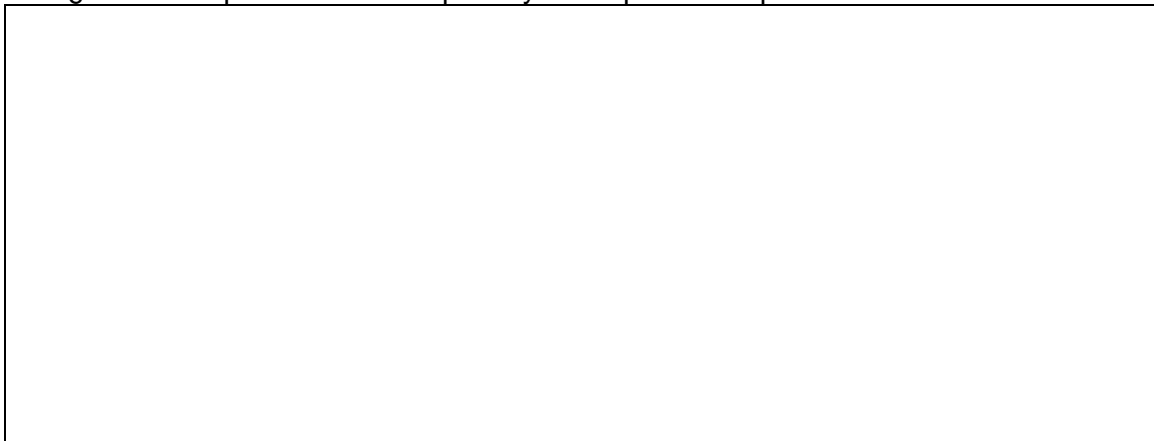
5.2 ¿Cuál es la probabilidad de que el foco defectuoso provenga de la caja que tiene 2% de focos defectuosos?



6. Una estación meteorológica suele acertar el 60% de las veces que pronostica día lluvioso. La probabilidad de que acierte su pronóstico dado de que el día no es lluvioso es 0,8. Se sabe que la probabilidad de que llueva un día cualquiera es 0,25. ¿Cuál es la probabilidad de que llueva si el pronóstico es correcto



7. En un almacén de una firma de fusibles se encuentra 80 cajas con 100 fusibles cada una, 20 cajas contienen fusibles producidos por una empresa A, 30 cajas contienen fusibles producidos por una compañía B, el resto de cajas contiene fusibles producidos por una compañía C. A produce el 3% de artículos defectuosos, B el 5% y C el 4% de artículos defectuosos. Si se selecciona una de estas cajas al azar, se toma uno de sus fusibles y se encuentra que es defectuoso. ¿Cuál es la probabilidad de que haya sido producida por B?



8. La compañía KANCIO está considerando comercializar una computadora. La probabilidad que la compañía tenga éxito es 0,8 si es que una firma competidora no introduce un producto similar en el mercado, en tanto que la posibilidad de éxito es sólo del 40% si la firma competidora introduce un producto similar. KANCIO estima que hay una probabilidad del 30% que la firma competidora comercialice el producto. Calcular:
- 8.1 La probabilidad que KANCIO no tenga éxito
- 8.2 Si KANCIO no tuvo éxito, ¿Cuál es la probabilidad que la competencia haya lanzado su producto?

## Autoevaluación

1. En una oficina hay dos computadoras A y B que trabajan de manera independiente. Si en un momento cualquiera la probabilidad de que la máquina B esté en mal estado es  $1/4$  y la probabilidad de que sólo la máquina A esté en mal estado es  $3/10$ . ¿Cuál es la probabilidad de que sólo la máquina B esté en malas condiciones?
2. Una persona está expuesta a un accidente en 10 ocasiones independientes. Si la probabilidad de que ocurra tal accidente es  $1/100$  cada vez, calcular la probabilidad de que:
  - 2.1 Un accidente ocurra en una o más ocasiones.
  - 2.2 3 accidentes ocurran, si uno de ellos ocurre en la décima ocasión.
3. Un experimento se realiza tantas veces en forma independiente hasta obtener el primer éxito. Suponga que en cada intento la probabilidad de que se tenga éxito es 0,95 si se siguen correctamente las instrucciones; y es de 0,20 si no se siguen correctamente las instrucciones. Calcular la probabilidad de alcanzar el éxito en tres intentos a lo más
  - 3.1 Si se siguen correctamente las instrucciones cada vez,
  - 3.2 Si no se siguen correctamente las instrucciones cada vez.
4. Suponga que un sistema funciona si al menos una de sus componentes funciona. Si las componentes trabajan independientemente y si la probabilidad que falle cada una es de 0,01, ¿cuántas componentes debería tener el sistema para que no falle con probabilidad de 0,9999?
5. Suponga que en cierta región del país la probabilidad de que un adulto mayor de 40 años tenga cáncer es 0,05, la probabilidad de que el diagnóstico sea correcto es 0,80, y de que sea errado es 0,20. si se elige al azar a una de esas personas, calcular, tenga realmente tal enfermedad.
  - 5.1 Se le diagnostique cáncer.
  - 5.2 Si se le diagnostica cáncer, tenga realmente tal enfermedad
6. Al contestar una pregunta de opción múltiple de 5 alternativas, donde sólo una es la respuesta correcta, un estudiante o bien conoce la respuesta o él responde al azar. La probabilidad de que conozca la respuesta es 0,6 y de que responda al azar es de 0,04.
  - 6.1 Halle la probabilidad que el estudiante conteste incorrectamente.
  - 6.2 Si contesta correctamente, calcular la probabilidad de que no conozca la respuesta.
7. Sólo el 60% de la mercadería que recibe un comerciante del fabricante A es de calidad excepcional, mientras que el 90% de la mercadería que recibe del fabricante B es de calidad excepcional. Sin embargo la capacidad de fabricación del fabricante B es limitada, y, por esta razón sólo el 30% de la mercadería le es permitido adquirir del fabricante B, el 70% la adquiere de A. Se inspecciona un embarque que acaba de llegar y se encuentra que es de calidad excepcional. ¿Cuál es la probabilidad de que provenga del fabricante A?

8. Un médico duda entre tres posibles enfermedades de un paciente designadas por E1, E2 y E3. Visto el estado general del paciente, al médico le parece que la E2 y la E3 son igualmente probables y que la E1 es tres veces más probable que la E2. Sin embargo ordena un análisis de sangre el que se sabe es positivo en el 10% de los casos cuando es la E1, 90% cuando es la E2 y el 60% de los casos cuando es la E3.
- 8.1 ¿Cuál es la probabilidad de que el análisis de sangre resulte positivo en el paciente que se está investigando?
- 8.2 Si el análisis de sangre resulta positivo, ¿seguirá siendo la E1 tres veces más probable que la E2 y que la E3?
9. El 100% de una población de electores se divide en tres estratos sociales excluyentes: baja, media y alta; de manera que la clase baja o media son el 90% del total, y la clase media o alta el 40% del total. De los primeros sondeos realizados para las próximas elecciones, se afirma que el porcentaje de electores que votarían por el candidato D puede ser: 30% de clase baja 50% de clase media 70% de clase alta.
- 9.1 Si se elige un elector al azar, ¿cuánto es la probabilidad de que:
- 9.1.1. vote por D,
- 9.1.1 Pertenezca al 10%. si vota por D
- 9-2 Si se escogen dos electores al azar, ¿qué probabilidad hay de que uno de ellos vote por D?
10. En un proceso de producción el porcentaje de objetos no defectuosos fabricados es 70% con probabilidad 0,35; 90% con probabilidad 0,25 y 60% con probabilidad 0,4. Si se selecciona al azar uno de tales objetos
- 10.1 Calcule la probabilidad de que sea defectuoso.
- 10.2 Si resulta no defectuoso, calcule la probabilidad de que sea de calidad del 90%:
11. Una máquina produce un tipo de objeto. Si la máquina está bien ajustada el 80% de los objetos producidos pasan el control de calidad, de otro modo sólo pasan el 60%. Se ha determinado que el 90% del tiempo la máquina está bien ajustado. De una caja que contiene 25 objetos de esa producción se escogen 3 al azar y a la vez para el control de calidad.
- 11.1 ¿Qué probabilidad hay que sólo 2 pasen el control de calidad?
- 11.2 Si sólo 2 pasen el control de calidad ¿qué probabilidad se tiene que haya sido producido cuando la máquina trabaja correctamente el 90% del tiempo?

## Para recordar

- Si al calcular la probabilidad de ocurrencia de un evento y esto depende de la ocurrencia de otros eventos, debemos utilizar el teorema de la probabilidad total.
- El uso del teorema de probabilidad total se facilita utilizando diagramas de árbol o cuadros de probabilidades.
- El teorema de Bayes nos permite calcular una probabilidad condicional en la que la ocurrencia del evento condicionante depende de la ocurrencia de otros eventos. Es por ello que previamente se debe calcular la probabilidad del evento condicionante utilizando el teorema de probabilidad total.
- Si queremos calcular la probabilidad de ocurrencia simultanea o sucesiva de una serie de eventos donde la ocurrencia de alguno de ellos no influye en la ocurrencia de cualquier otro (no existe dependencia) debemos utilizar el teorema de ocurrencia simultanea de eventos independientes.





# DISTRIBUCIÓN BINOMIAL

## TEMAS

- Distribuciones de probabilidad discreta. Distribución Binomial

## OBJETIVOS ESPECÍFICOS

- Distinguir cuando una variable aleatoria se distribuye binomialmente y calcular las probabilidades asociadas a dicha variable

## CONTENIDOS

- Distribución de probabilidad
- Distribución Binomial

## ACTIVIDADES

- Reconocen los problemas de distribución binomial y aplicar su formula para su desarrollo

## Distribuciones de Probabilidad

**Prueba de Bernoulli:** Un experimento aleatorio es considerado una prueba de Bernoulli si cumple lo siguiente:

1. En cada prueba sólo hay dos resultados posibles denominados arbitrariamente éxito o fracaso
2. La probabilidad de éxito ( $\pi$ ) se mantiene constante prueba a prueba que hay independencia de prueba a prueba.

**Distribución Binomial:** es un experimento aleatorio constituido por "n" pruebas de Bernoulli y para la cual se define la variable aleatoria discreta X: Número de "éxitos" en "n" pruebas.

Notación:  $X \sim B(n, \pi)$

Rango:  $X \in \{0, 1, 2, 3, 4, \dots, n\}$

Función de probabilidad:

$$f(x) = C_x^n \pi^x (1 - \pi)^{n-x}$$

Valor Esperado:  $E(x) = n \cdot \pi$

Varianza:  $V(x) = n \cdot \pi \cdot (1 - \pi)$

## Actividades

1. Si X denota el número de caras en un solo lanzamientos de 4 monedas.

Calcular:

- 1.1  $P(X = 3)$
- 1.2  $P(X < 2)$
- 1.3  $P(X \geq 2)$
- 1.4  $P(1 < X \leq 3)$

2. Si se lanzan 6 monedas al mismo tiempo, hallar la probabilidad de obtener:

- 2.1 Dos ó más caras
- 2.2 Menos de 4 caras
- 2.3 Calcule el valor esperado y la varianza de X

3. Se sabe que el 40% de las personas que se presentan a un empleo son altamente capacitadas en el manejo de Visual Basic. Si 25 personas se presentan a un empleo:
- 3.1 Defina la variable aleatoria y su función de probabilidad
  - 3.2 ¿Cuál es la probabilidad que exactamente 10 sean altamente capacitadas en VB?
  - 3.3 ¿Cuál es la probabilidad que por lo menos 2 sean altamente capacitadas en el manejo de Visual Basic?

4. Una computadora no funciona el 20% de las veces que se prende. ¿Cuál es la probabilidad que en 10 intentos:
- 4.1 No funcione exactamente 5 veces
  - 4.2 No funcione por lo menos dos veces
  - 4.3 Funcione exactamente 4 veces

## Auto evaluación

1. Una máquina produce cierto tipo de piezas, de las cuales un promedio de 5% es defectuosa. En una muestra aleatoria de cinco piezas. ¿Cuál es la probabilidad de obtener:
  - 1.1 Exactamente una pieza defectuosa
  - 1.2 Por lo menos una pieza defectuosa
2. En una población de drosophila, el 20% tienen mutación de alas. Si se escogen 6 moscas aleatoriamente de la población.
  - 2.1 ¿Cuál es la probabilidad que dos tengan mutación?
  - 2.2 ¿Al menos uno tenga mutación?
  - 2.3 ¿Qué menos de 5 tengan mutación?
  - 2.4 ¿Cuál es el número esperado de moscas con mutación de alas?
3. Un tratamiento para cierta enfermedad produce una cura en 75% de los casos. Se seleccionan 6 pacientes aleatoriamente. ¿Cuál es la probabilidad de que:
  - 3.1 Todos estén curados.
  - 3.2 Ninguno esté curado.
  - 3.3 Cuatro estén curados
  - 3.4 Al menos cuatro estén curados
4. El 90% de los tubos de ensayo soportan una temperatura mayor que 80°C; suponga que 10 de estos tubos se someten a una prueba a temperaturas mayores de 80°C. Determine la probabilidad que 3 de estos tubos quede inutilizables.
5. La probabilidad de fallar durante el vuelo para cada uno de los seis motores de un avión es 0,0005. Suponiendo que los seis motores trabajan independientemente, determine la probabilidad que en un vuelo determinado:
  - 5.1 No ocurra ninguna falla del motor.
  - 5.2 No ocurra más de una falla.
  - 5.3 Ocurra exactamente dos fallas.
6. Suponga que los motores de un avión de cierta marca, que operan independientemente, tienen una probabilidad de falla de 0,1. Suponga que un avión efectúa un vuelo exitoso si al menos la mitad de sus motores operan normalmente, determine cuál avión, uno con cuatro y otro con seis motores, tiene mayor probabilidad de efectuar un vuelo exitoso.
7. Cierta tubo de televisión tiene una probabilidad de 0,3 de funcionar más de 400 horas. Se prueban 15 tubos
  - 7.1 Hallar la posibilidad que exactamente 0, 4, 9 de ellos funcione más de 400 horas.
  - 7.2 Cuántos tubos espera encontrar que funcionen por lo menos 400 horas
  - 7.3Cuál es el número de tubos más probable que funcionen por lo menos 400 horas.
8. Suponga que la máquina A produce el doble de artículos que la máquina B. Se sabe que el 6% de los artículos que produce la máquina A son defectuosos, mientras que el 3% de los artículos producidos por la máquina B son defectuosos. Suponga que se junta la producción diaria de estas máquinas y se toma una

muestra aleatoria de 10 artículos. Calcular la probabilidad de obtener 3 artículos defectuosos.

9. La probabilidad de falla de los motores de un avión que tiene en total seis motores es 0,0005. Determine la probabilidad que en un vuelo determinado:
  - 9.1 No ocurra ninguna falla de motor.
  - 9.2 Ocurra más de una falla.
10. De todas las personas que han pasado una prueba de aptitud a un trabajo en los últimos años, 90% han tenido éxito desempeñando ese trabajo, ¿Cuál es la probabilidad de que cinco de seis solicitantes que han pasado la prueba recientemente tengan éxito en el trabajo? ¿Cuál es la probabilidad que por lo menos cinco tengan éxito?
11. Suponga que el 20% de las solicitudes de préstamo en un banco son aceptadas. Si se elige al azar a 8 personas que solicitaron el préstamo, ¿Cuál es la probabilidad que
  - 11.1 A 2 de ellas se le acepte el préstamo
  - 11.2 A por lo menos 5 se le acepte el préstamo
  - 11.3 Se le acepte el préstamo a más de 6 personas si se sabe que se le ha aceptado a por lo menos 2
12. Se ha elaborado un examen de selección múltiple consistente en 10 preguntas. Hay 5 respuestas posibles para cada pregunta. Suponga que ninguno de los estudiantes que va a rendir el test concurre a clase o que no estudio para el examen. El profesor que toma el examen ha establecido que para aprobar debe contestar correctamente al menos 6 preguntas. ¿Cuál es la probabilidad de aprobar.

## Para recordar.

- En una distribución binomial se debe tener en cuenta que la variable a estudiar debe ser discreta.
- Para un proceso binomial se debe tener en cuenta la muestra a estudiar y la probabilidad de éxito.



# DISTRIBUCIÓN DE POISSON

## TEMAS

- Distribución de Poisson..

## OBJETIVOS ESPECÍFICOS

- Distinguir cuando una variable aleatoria se distribuye según Poisson y calcular las probabilidades asociadas a dicha variable..

## CONTENIDOS

- Distribución de Poisson..

## ACTIVIDADES

- Reconocen los problemas de distribución Poisson y aplicar su formula para su desarrollo.

## Distribuciones de Probabilidad

**Distribución de Poisson:** es un experimento aleatorio para el cuál se define la variable aleatoria discreta  $X$ : Número de “éxitos” en un intervalo (tiempo, espacio, área, etc.)

Notación:  $X \sim P(\lambda)$

Donde  $\lambda$  es un valor promedio conocido para el intervalo en que está definida la variable aleatoria.

Función de probabilidad :

$$f(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

Valor Esperado:  $E(x) = \lambda$

Varianza:  $V(x) = \lambda$

### Actividades:

1. Una de cada cinco veces que se llegó a un puente levadizo, éste se encontraba levantado por lo que se tuvo que detener el auto para luego continuar el camino.
  - 1.1 Calcular la probabilidad que en las siguientes siete aproximaciones al puente, este no se halle levantado.
  - 1.2 ¿Cuál es el número de veces que se espera encontrar el puente levantado?

2. Los accidentes de trabajo, que se producen por semana en una fabrica, siguen la ley de Poisson, de tal manera que la probabilidad que haya 5 accidentes es 16/15 de que haya 2.
  - 2.1 Halle el parámetro de la distribución de Poisson.
  - 2.2 La probabilidad que no haya accidentes en tres semanas.



3. Una cajera de un supermercado puede atender en promedio a dos clientes por minuto. Calcule la probabilidad para que:
- 3.1 Pueda atender a 5 clientes en el lapso de dos minutos.
- 3.2 Pueda atender por lo menos a dos clientes en el lapso de un minuto.

4. Suponga que cierto alimento produce una reacción alérgica en un 0,01% de una población grande. Si 100 000 personas comen este alimento diario en promedio.
- 4.1 Cuál es la probabilidad que 5 de ellos sean, alérgicos a este alimento.
- 4.2 Cuál es el número esperado de personas con reacción alérgica.

5. Cierta alimento produce una reacción alérgica en un 0,01% de una población grande. Si 100 000 personas comen este alimento diario en promedio.
- 5.1 ¿Cuál es el número esperado de personas con reacción alérgica?
- 5.2 ¿Cuál es la función de probabilidad del número de personas en este grupo de 100 000 son alérgicos a este alimento?

# Autoevaluación

1. En promedio, cada hora cinco personas realizan transacciones en el mostrador de servicios especiales. Suponiendo que la llegada de estas personas es independiente e igualmente probable:
  - 1.1 ¿Cuál es la probabilidad de que más de 3 personas efectúen transacciones en una hora específica?
  - 1.2 Si en el transcurso de una hora, más de tres personas ya han efectuado transacciones, ¿Cuál es la probabilidad que sean menos de seis las personas que efectúan transacciones en dicha hora?
2. En una intersección cruzan, en promedio, 5 autos en 10 segundos. El semáforo de dicha intersección mantiene la luz verde durante 15 segundos. Si en un momento determinado llegan al semáforo 8 autos y justo en ese instante la luz cambia a verde, ¿Cuál es la probabilidad que dos autos no logren cruzar la intersección?
3. Un sistema puede atender en promedio a 2 operadores simultáneos en una hora. Sin embargo el sistema colapsa cuando es requerido por más de 3 operadores simultáneamente en una hora.
  - 3.1 ¿Cuál es la probabilidad que el sistema colapse?
  - 3.2 Si se define X: Número de veces que el sistema colapsa en 24 horas, determine la distribución de probabilidades de X.
4. Se ha observado que el promedio de ventas del producto “A” en una empresa es de 3 unidades por hora. Si se supone que las ventas son independientes una de otra, y si X representa el número de ventas cada 20 minutos ¿Cuál es la probabilidad que;
  - 4.1 En el intervalo de 20 minutos no se realice venta alguna
  - 4.2 Se realice al menos 2 ventas, en el intervalo de 20 minutos
5. En una compañía el sistema queda suspendido a razón de 2 veces por mes. Determine la probabilidad que:
  - 5.1 En dos meses, el sistema quede suspendido por lo menos 4 veces.
  - 5.2 En dos meses el sistema quede suspendido más de 5 veces, si se sabe que luego de un mes el sistema ya quedó suspendido una vez.
6. Un cajero automático esta programado para atender en promedio a dos movimientos de los clientes por minuto. Calcule la probabilidad que:
  - 6.1 Pueda atender a 4 movimientos de los clientes en el lapso de dos minutos.
  - 6.2 Pueda atender por lo menos a tres movimientos de los clientes en el lapso de un minuto.
7. El número de casos admitidos de emergencia en cierto hospital es 1 hora es una variable aleatoria con distribución de Poisson con  $\lambda = 3$ . Determine que en cierta hora:
  - 7.1 Ningún caso de emergencia sea admitido.
  - 7.2 Más de tres casos de emergencia sean admitidos

8. Suponga que cierta enfermedad rara afecta al 0,1% de la población grande. Se escogen aleatoriamente a 5 000 personas de esta población y son sometidos a un examen para detectar la enfermedad.
  - 8.1 ¿Cuál es el número esperado de personas con dicha enfermedad?
  - 8.2 ¿Cuál es la probabilidad que exactamente 10 personas queden afectadas por la enfermedad?
9. Suponga que un libro de 585 páginas contiene 43 errores tipográficos. Si estos errores se distribuyen aleatoriamente a través del libro. ¿Cuál es la probabilidad que 10 páginas, seleccionadas al azar, no contengan errores?
10. La probabilidad de que una persona sufra una reacción alérgica a un determinado medicamento es 0,001. Determine la probabilidad de que un total de 2 000 personas que han tomado el medicamento,
  - 10.1 Exactamente 3 tengan reacción alérgica.
  - 10.2 Más de dos personas tengan reacción alérgica
  - 10.3 Halle el número esperado de personas con reacciones alérgicas y su varianza.
11. Suponga que la probabilidad de que un motor falle en un vuelo de rutina entre dos ciudades es 0,005. Use la aproximación de Poisson a la distribución Binomial para encontrar aproximadamente la probabilidad de que:
  - 11.1 Por lo menos una falla en 1 000 vuelos.
  - 11.2 Por lo menos dos fallas en 1 000 vuelos.
12. La probabilidad de que se haga una soldadura defectuosa en una conexión dada es  $10^{-4}$ . Considere un sistema de  $5 \times 10^4$  conexiones soldadas independientemente.
  - 12.1 ¿Cuál es la distribución de probabilidad del número de uniones defectuosas en el sistema? ¿Cuáles son los parámetros?
  - 12.2 Utilice, una aproximación, a la verdadera distribución, para calcular la probabilidad de que no se presenten defectos en el sistema.

## Para recordar.

- En una distribución de Poisson se debe tener en cuenta que la variable a estudiar debe ser discreta.
- Para un proceso de Poisson se debe tener en cuenta la ocurrencia de eventos discretos en un intervalo continuo.



# DISTRIBUCIÓN NORMAL

## TEMAS

- Distribución de probabilidad continua.
- Distribución Normal. Características
- Uso de la tabla de distribución Normal Estándar.

## OBJETIVOS ESPECÍFICOS

- Conocer la distribución Normal, sus propiedades y su importancia.
- Usar la tabla de la distribución Normal Estándar para calcular probabilidades de variables normalmente distribuidas.

## CONTENIDOS

- Función de probabilidad.
- Distribución normal.
- Distribución normal estándar.
- Distribución Normal

## ACTIVIDADES

- Reconocen los problemas de distribución Normal y aplicar su formula para su desarrollo.
- Manejo adecuado de la tabla de distribución Normal y aplicarla como herramienta de desarrollo.

## DISTRIBUCION DE PROBABILIDAD CONTINUA

### Función de probabilidad:

Se define a  $f(x)$  como la función de probabilidad o función de densidad de  $X$  si cumple con las siguientes condiciones:

$$1.1. \quad f(x) \geq 0, \quad \text{Para todo } x_i \in R_x$$

$$1.2. \quad \int_{-\infty}^{+\infty} f(x) = 1$$

$$1.3. \quad \int_a^b f(x) = P(a \leq x \leq b)$$

Ejemplo: Si la variable aleatoria  $x$  tiene como función de densidad:

$$f(x) = 2, \quad \text{si: } 0 < x < 1/2$$

$$0, \quad \text{en otro caso}$$

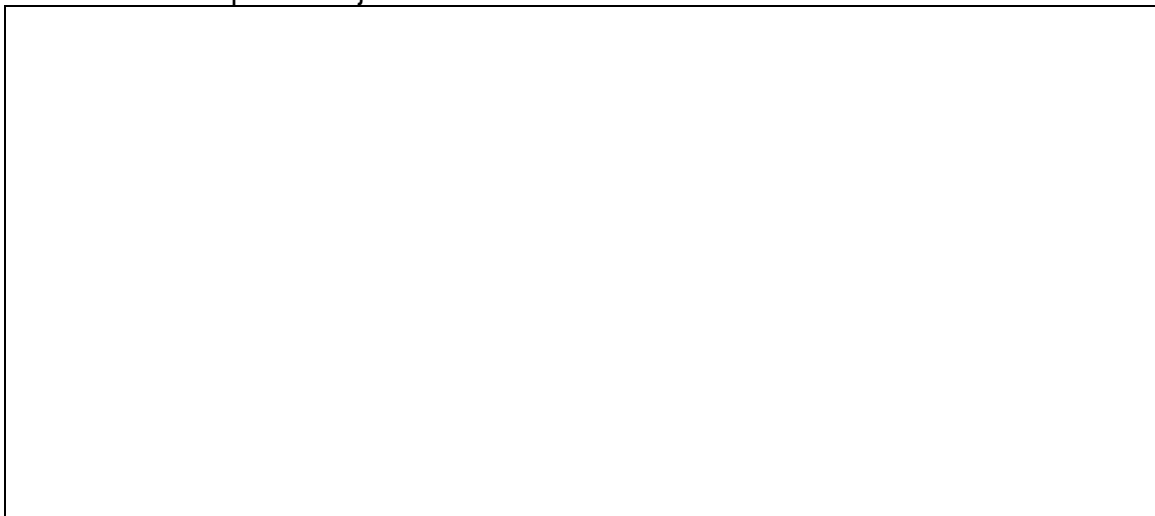
### Distribución Normal:

Una variable aleatoria  $X$  que tiene la función de densidad

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2}$$

Se dice que sigue una distribución normal con media  $\mu$  y desviación estándar  $\sigma$ . ( $\mu$  y  $\sigma$  poblacionales).

La gráfica de  $f(x)$  tiene forma acampanada, es simétrica respecto de la recta  $x = \mu$  y es asintótica respecto al eje horizontal.



Notación:  $X \sim N(\mu, \sigma^2)$

**Distribución Normal Estándar:**

Una variable aleatoria continua  $Z$  tiene distribución normal estándar si sigue una distribución normal con  $\mu = 0$  y  $\sigma = 1$  ( $Z \sim N(0, 1)$ )

Esta distribución nos permite calcular la probabilidad de que la variable aleatoria  $Z$  tome un valor determinado mediante el uso de tablas previamente desarrolladas.

Uso de tablas: Si  $Z \sim N(0, 1)$ , determine:

- |                         |                          |
|-------------------------|--------------------------|
| a) $P(Z < 1.2)$         | b) $P(Z < -2.21)$        |
| c) $P(0.89 < Z < 1.52)$ | d) $P(-1.15 < Z < 2.02)$ |
| e) $P(Z > 0.78)$        | f) $P(Z > -1.33)$        |
| g) $P( z  < 1.56)$      | h) $P( z  > 0.63)$       |



La distribución normal estándar es utilizada para determinar la probabilidad de ocurrencia de una variable aleatoria que sigue una distribución normal cualquiera. Para ello, previamente, se debe estandarizar (transformar) la variable original, mediante la siguiente transformación:

$$Z = \frac{X - \mu}{\sigma}$$

Ejemplo: Si  $X \sim N(16, 25)$ , determine:

- a)  $P(x < 19)$
- b)  $P(x < 11)$
- c)  $P(13 < x < 17.6)$
- d)  $P(x > 17.75)$





## ACTIVIDADES:

1. Los pesos de las impresoras marca ZZ se distribuyen normalmente con media 1,5 kg. Y desviación estándar 0,3 kg. Determine:
  - 1.1 La probabilidad que una impresora elegida al azar pese menos de 1,3 Kg.
  - 1.2 El porcentaje de impresoras que pesan entre 1,4 y 1,65 Kg.
  - 1.3 Si la producción total de impresoras es de 5 000, ¿Cuántas pesan más de 1,75 Kg.?

2. El tiempo que demora “cargar” un sistema se distribuye normalmente con media 20 seg. Y desviación estándar de 2,5 seg. Determine:
  - 2.1 La probabilidad que el sistema se demore en cargar menos de 18 seg.
  - 2.2 El porcentaje de veces que el sistema se demora en cargar más de 21 seg.

3. Del aeropuerto al centro de la ciudad se puede llegar mediante dos caminos. Si se viaja por el camino A, el tiempo está distribuido normalmente con media 27 minutos y desviación estándar de 5 minutos. Si se viaja por el camino B el tiempo está distribuido normalmente con media 25 minutos y desviación estándar de 12 minutos, ¿Por cuál camino es preferible ir al centro de la ciudad si se dispone de 30 minutos?

4. Usted debe decidir acerca de qué sistema se debe utilizar en una gran empresa. El sistema XX tiene un rendimiento promedio de 110 000 horas con una desviación estándar de 25000 horas, en cambio, el sistema YY tiene un rendimiento promedio de 120 000 horas con una desviación estándar de 40 000 horas. Basándose en criterios netamente probabilísticos ¿Cuál sistema elige, si dicho sistema debe tener un rendimiento mínimo de 125 000 horas, ¿Cuál sería su decisión? Justifique su respuesta.

5. Se ha encontrado que el tiempo de servicio que se requiere por persona en una caja bancaria tiene una distribución aproximadamente normal con media de 130 segundos y desviación estándar de 45 segundos.
- 5.1 Calcule la probabilidad de que se requiera menos de 100 segundos
- 5.3 ¿Cuál es el tiempo mínimo que utilizan el 20% de las personas que más se demora.

6. Supongamos que las ventas mensuales obtenidas por un grupo de vendedores de una tienda se distribuyen normalmente con media 15 000 soles y desviación estándar 750 soles. Si al 15% de los que más venden se les bonificará con 200 soles y al 25% de los que siguen en mérito por el monto vendido se les bonificará con 100 soles, ¿Cuánto desembolsará la tienda por concepto de bonificaciones, si esta cuenta con 2 500 vendedores?

7. Se está construyendo un grupo de 100 casas en la Urbanización San Borja. El material empleado en las redes de desagüe es tal que el 9,512% de las tuberías de desagüe tiene periodos de duración que exceden los 15 años y que el 62,556% tienen periodos de duración que exceden los 9 años. Considerando que la distribución de probabilidad de los periodos de duración de éstas tuberías es normal, determínese la media y la varianza de esta distribución.

8. Un supermercado almacena 30 kilogramos de queso fundido cada semana. Si la demanda semanal de queso fundido está normalmente distribuida con media 24 kilogramos y desviación típica 5 kilogramos. Determinar la probabilidad que el supermercado agote los quesos fundidos durante una semana seleccionada al azar.

9. El bar “Un par más” ha instalado una máquina automática para la venta de cerveza. La máquina puede regularse de modo que la cantidad media de cerveza por vaso sea la que se desea; sin embargo, en cualquier caso esta cantidad tendrá una distribución normal con una desviación estándar de 5,9 mililitros.

9.1 Si el nivel se ajusta a 304,6 mililitros, ¿Qué porcentaje de los vasos contendrán menos de 295,7 mililitros?

9.2 ¿A qué nivel medio debe ajustarse la máquina para que sólo el 2,28% de los vasos contengan menos de 295,7 mililitros?

9.3 ¿A qué nivel medio debe ajustarse la máquina para que el 84,13% de los vasos contenga menos de 313,6 mililitros?

10. La presión sanguínea media en hombres de 20 a 25 años de edad es 123 unidades con desviación típica de 13,7 unidades. Si se selecciona al azar uno de estos hombres, calcule la probabilidad que su presión sanguínea esté comprendida entre 120 y 128 unidades. Suponga una distribución normal.

## Autoevaluación

1. Los sueldos de los trabajadores de una empresa transnacional se distribuyen normalmente con una media de 2 240 soles y una desviación estándar de 510 soles. ¿Cuántos trabajadores tiene la empresa si hay 84 personas que ganan entre 2 100 soles y 2 500 soles?
2. Una máquina automática para el llenado de paquetes de café puede regularse de modo que la cantidad media de arroz llenado sea la que se desee. Si la cantidad de café depositada se distribuye normalmente con una varianza igual a 100 gr<sup>2</sup> ¿Cuál debe ser el valor de la regulación media de modo que sólo el 1% de los paquetes tengan un peso neto inferior a 990 gr?
3. El tiempo que se requiere para reparar cierto tipo de transmisión automotriz en un taller mecánico tiene distribución normal con media 45 minutos y desviación estándar de 8 minutos. El gerente de servicio planea hacer que se inicie la reparación de la transmisión de los automóviles de los clientes, diez minutos después de que se recibe el vehículo y le dice al cliente que el vehículo estará listo en una hora. ¿Cuál es la probabilidad que el gerente esté equivocado?
4. Una fábrica de neumáticos hace una prueba para medir la duración de sus llantas y verificó que la duración de las llantas tenía distribución normal, con media 48 000 Km., y desviación estándar de 2 000 Km.
  - 4.1 Calcular la probabilidad que una llanta escogida al azar dure entre 45 000 Km. y 50 000 Km.
  - 4.2 El fabricante quiere garantizar que la llanta durará por lo menos “N” Km., ¿Cuál deberá ser el valor de “N” para que la probabilidad de que se equivoque sea menor o igual a 5.05%?
  - 4.3 Las llantas se venden asegurando que duran más de 45 000 Km. ¿Cuál es la probabilidad que de 5 llantas vendidas, exactamente tres tengan una duración menor a la asegurada?
5. Los tiempos de vida de las baterías para automóvil marca WV tiene distribución normal con media de 6 750 h. y una varianza de 250 000 h<sup>2</sup>. Según el tiempo de vida de las baterías, éstas son clasificadas de la manera siguiente: de clase A, si el tiempo de vida es menor de 6,000 h.; de clase B, si el tiempo de vida está comprendido entre las 6 000 y 7 400 h.; y de clase C, si el tiempo de vida es mayor de 7 400 h.
  - 5.1 Determine el porcentaje de baterías de cada clase que son fabricadas.
  - 5.2 Si las baterías se transportan en cajas de 10 unidades, ¿qué porcentaje de cajas tendrá más de una batería clase A?
6. Las estaturas de 1 000 estudiantes están distribuidas normalmente con una media de 174,5 centímetros y una desviación estándar de 6,9 centímetros. Suponiendo que las estaturas se redondean al medio centímetros más cercano, ¿cuántos de estos estudiantes se espera que tengan estaturas?
  - 6.1 Menores de 160cm.
  - 6.2 Entre 171,5 y 182cm. Inclusive
  - 6.3 Mayores o iguales a 188 cm.

6.4 Iguales a 175cm.

7. El número de días entre la facturación y el paso de las cuentas corrientes de crédito en una tienda de departamentos grande tiene una distribución aproximadamente normal con una media de 18 días y una desviación estándar de 4 días. ¿Qué proporción de las facturas será pagada?
  - 7.1 Entre 12 y 18 días
  - 7.2 Entre 20 y 23 días
  - 7.3 En menos de 8 días
  - 7.4 ¿Dentro de cuántos días estará pagado el 99,5% de las facturas?
  - 7.5 Entre cuales dos valores simétricamente distribuidos alrededor de la media recaerá el 98% de las facturas
8. La medición del diámetro de contacto de la rosca de una unión se distribuye normalmente con media 0,4008 pulgadas y una desviación estándar 0,003 pulgadas. ¿Cuál es la probabilidad que una rosca tenga medidas entre 0,40002 y 0,40013 pulgadas?

## Para recordar.

- En una distribución Normal se debe tener en cuenta que la variable a estudiar debe ser continua.
- Para un proceso Normal se debe tener en cuenta la media promedio y la desviación estándar.
- Para el mejor manejo de la tabla, el valor encontrado debe aproximarse a la milésima.







# DISTRIBUCIÓN NORMAL

## TEMAS

- Teorema del límite Central.
- Aproximación de la distribución binomial mediante la distribución normal

## OBJETIVOS ESPECÍFICOS

- Comprende la importancia del Teorema del Límite Central y sus aplicaciones.

## CONTENIDOS

- Teorema del Límite Central.
- Aproximación de la distribución binomial mediante la distribución normal

## ACTIVIDADES

- Reconocen los problemas de distribución Normal a partir del teorema del Límite Central

## TEOREMA DE LÍMITE CENTRAL

Se dice que  $n$  variables aleatorias  $X_1, X_2, X_3, \dots, X_n$  tienen la misma distribución de probabilidad con media  $\mu$  y varianza  $\sigma^2$ , si tienen la misma función de probabilidad en el caso de variable discreta, o una misma función de densidad en el caso de variable continua.

Si :  $X_1, X_2, X_3, \dots, X_n$  son  $n$  variables aleatorias independientes igualmente distribuidas con media  $\mu$  y varianza  $\sigma^2$ , entonces la distribución de la variable aleatoria:

$$Z_n = \frac{(X_1 + X_2 + X_3 \dots + X_n) - n\mu}{\sigma \sqrt{n}}$$

## ACTIVIDADES

1. Una linterna grande es alimentada por cinco baterías. Suponga que la vida de una batería está normalmente distribuido con media de 120 horas y desviación estándar de 10 horas. La linterna dejará de funcionar si se agota una o más de sus baterías. Suponiendo que las vidas de las baterías son independientes. ¿Cuál es la probabilidad que la linterna funcione mas de 100 horas?

2. Los pesos de los paquetes recibidos en un departamento de almacenamiento tienen una media de 300 libras y una desviación típica de 50 libras. ¿Cuál es la probabilidad de que el peso de 25 paquetes recibidos al azar y cargados en un ascensor supere el límite de seguridad del ascensor, que es de 8 200 libras?

3. El peso de las impresoras marca XX modelo G 1 560 se distribuye normalmente con media 1,2 Kg., y con desviación estándar de 0,6 Kg. Estas impresoras se colocarán en cajas que pueden soportar como máximo 16 Kg. ¿Cuál es la probabilidad que 11 impresoras colocadas en una de las cajas supere la capacidad de carga de la misma?

4. Un camión de reparto transporta cajones cargados de artículos varios. Si el peso de cada cajón se distribuye normalmente con media 50 lb. y una desviación estándar 5 lb. ¿Cuántos cajones pueden ser transportados en el camión de tal forma que la probabilidad que la carga total exceda a 1 tonelada sea sólo 0,10?

## Aproximación de la distribución binomial mediante la distribución normal

Una variable aleatoria se distribuye binomialmente con parámetros  $n$  y  $\pi$ . Si se define la variable  $X = X_1 + X_2 + X_3 + \dots + X_n$ , donde cada  $X_i$  es una variable aleatoria de Bernoulli con valor esperado  $\pi$  y varianza  $\pi(1-\pi)$ . Si  $n$  es grande y  $\pi$  tiende a 0.5 se puede realizar aproximar a la normal de la siguiente forma:

$$X \approx N [n\pi, n\pi(1-\pi)]$$

Nota: La aproximación es mejor si:  $n\pi > 5$

Observación: Al usar la normal como aproximación de la binomial, se aproxima la distribución de una variable aleatoria discreta con la de una variable aleatoria continua. Por ello debe hacerse la corrección por continuidad.

## ACTIVIDADES

1. Una firma comercializa sus productos vía Internet a una lista de 100 000 clientes potenciales, para decidir acerca de la comercialización de un nuevo producto. La firma selecciona una muestra de 100 personas para ofrecerles dicho artículo. Si 30 o más de estos clientes están dispuestos a adquirirlo, se procederá a su comercialización.
  - 1.1 Cual es la probabilidad que comercialice el artículo, si en realidad sólo el 20% de todos los clientes lo comprasen.
  - 1.2 Cual es la probabilidad que comercialice el artículo, si en realidad sólo el 36% de todos los clientes lo comprasen.

2. En un almacén se tienen 300 fusibles, y por experiencia se sabe que hay un 2% de defectuosos en la producción de fusibles.
  - 2.1 Calcular la probabilidad que exista exactamente un defectuoso.
  - 2.2 Calcular la probabilidad que no haya mas de 5 defectuosos.

3. En Una población grande de moscas, el 25% tienen mutación de alas. Una muestra de 300 insectos es escogida al azar. Calcule la probabilidad que más de 60 pero no mas que 90 insectos de la muestra tengan mutación de alas.

4. Se sabe que cierto virus ha invadido una escuela y ataca a la mitad de los estudiantes. Se toma una muestra aleatoria de 200 alumnos:
- 4.1 Calcule la probabilidad que en dicha muestra el 49% sea atacada por el virus.
  - 4.2 Calcular la probabilidad que en la misma muestra ninguno presente síntomas del virus.

## Autoevaluación

1. Las pastillas metálicas cilíndricas que se utilizan en un reactor se fabrican en serie y puede suponerse que sus longitudes siguen una distribución normal con media de 0,290 cm. Y desviación estándar 0,016 cm. Nueve de estas pastillas deben ajustarse, extremo con extremo, en un recipiente que ocupa una longitud no mayor de 2,670 cm. Si las nueve pastillas se ensamblan al azar, ¿Qué proporción de estos no se ajustará en el espacio requerido?
2. En una de las etapas de un proceso de ensamble un tapón cilíndrico tiene que ajustarse a una abertura circular seleccionando cada elemento al azar en un suministro continuo. Los diámetros del tapón Mm. , son  $N(24,6; 0,03^2)$  . Si para que el ajuste sea satisfactorio se requiere un claro en el diámetro de al menos 24,2 Mm. y a lo más de 29,8 Mm., ¿En qué proporción de los casos el ajuste no será satisfactorio?
3. Un tipo de censor tiene una vida de 15 000 horas con una desviación estándar de 1 000 horas. Tres de estos sensores se conectan en un dispositivo de rector de incendios de manera que cuando uno falla el otro lo sustituye. Suponiendo que la durabilidad tiene una distribución normal, Cuál es la probabilidad que el dispositivo:
  - 2.1 Funcione durante 40 000 horas.
  - 2.2 Falle antes de 39 000 horas
4. Suponga que los pesos de los pasajeros que viajan por aire en los vuelos establecidos que parten de un aeropuerto grande siguen una distribución normal con media de 78kg. y desviación estándar 10kg. Encuentre los límites (simétricos alrededor de la media) de tal manera que el 95% de los pasajeros tengan un peso límite dentro de estos valores. ¿Cuál es la probabilidad de que peso total de una muestra al azar de 100 pasajeros exceda a 8 000 Kg.?
5. Dos marcas de focos “Económica” y “vida eterna” tienen durabilidad (en horas) que son  $N(1\,400, 200^2)$  y  $N(2\,000, 250^2)$  respectivamente. Si se prueba la vida de uno de los focos correspondientes a cada una de las marcas, ¿Cuál es la probabilidad que la “económica” dure más?
6. Si el 10% de los circuitos integrados de radio se malogran antes que la garantía haya espirado. Cual es la probabilidad, que al vender 100 de estos integrados.
  - 6.1 Se vea obligado a sustituir por lo menos a 20 de ellos.
  - 6.2 Sustituye 5 y no más de 15 tubos.
7. Con base en la experiencia pasada, el 40% de todos los clientes de cierto grifo pagan sus compras con tarjetas de crédito. Si se selecciona una muestra aleatoria de 200 clientes, ¿Cuál es la probabilidad que?
  - 7.1 Cuando menos 75 paguen con tarjetas de crédito.
  - 7.2 No más de 70 paguen con tarjetas de crédito.
  - 7.3 Entre 70 y 75 clientes, inclusive, paguen con tarjeta de crédito.
8. Si una muestra de 100 tiene 3 o menos artículos defectuosos se acepta el lote de 100. Si se supone que la probabilidad de producir artículos buenos del proceso de producción es de 0,90. ¿Cuál es la probabilidad de que se acepte el lote?
9. Una compañía planea recibir de vuelta alrededor del 30% de los cupones especiales de premio que planea enviar por correo para un programa de promoción de ventas.

Si se enviaran 500 cupones, ¿cuál es la probabilidad de que se reciban de vuelta más de 165 respuestas?

10. Suponga que un 10% de las llantas de un fabricante tienen defectos en la superficie, y que los embarca en lotes de 100.
  - 10.1 ¿Cuál es la probabilidad aproximada de que un lote contenga 8 ó menos llantas con defectos en su superficie?
  - 10.2 Un comprador mayorista recibe 500 lotes. ¿Cuál es la probabilidad aproximada de que al menos 140 lotes contengan 8 ó menos llantas con defectos en la superficie cada uno?

## Para recordar.

- En el uso de limite central se debe tener en cuenta que las  $n$  variables aleatorias independientes a estudiar deben ser continuas.
- Para un proceso Normal se debe tener en cuenta la media promedio y la desviación estándar común y finitas.
- En el uso de la aproximación de la distribución binomial a la normal se realizara cuando la muestra sea grande.
- Para un proceso de aproximación debe tenerse en cuenta los parámetros tomados para la aproximación.





# **DISTRIBUCIÓN MUESTRAL DE MEDIAS Y DE PROPORCIONES**

## **TEMAS**

- Distribución muestral de medias y proporciones

## **OBJETIVOS ESPECÍFICOS**

- Elaborar una distribución de medias y proporciones muestrales.
- Cálculo de probabilidades asociadas a esta distribución.

## **CONTENIDOS**

- Distribución muestral de medias
- Distribución muestral de proporciones

## **ACTIVIDADES**

- Aproximan a la distribución normal cuando  $n$  tiende al infinito.

## DISTRIBUCIÓN MUESTRALES PARA MEDIAS Y PROPORCIONES

El Teorema del Límite Central es un teorema fundamental en la estadística. Su aplicación permite determinar la distribución de la media y de proporción muestral

### Distribución muestral de medias

Si una muestra aleatoria de tamaño "n" es extraída de una población con distribución dada, con media  $\mu$  y varianza  $\sigma^2$ , entonces la distribución de la v.a. media muestral ( $\bar{X}$ ) se aproxima cada vez más a una distribución normal cuando "n" tiende a infinito.

$$\bar{X} \approx N(\mu_{\bar{x}}, \sigma_{\bar{x}}^2)$$

Donde:  $\mu_{\bar{x}} = \mu$  y  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

### Distribución muestral de proporciones

Si una muestra aleatoria de tamaño "n" es extraída de una población con distribución dada, con proporción poblacional ( $\pi$ ), entonces la distribución de la v.a. proporción muestral (p) se aproxima cada vez más a una distribución normal cuando "n" tiende a infinito.

$$p \approx N(\mu_p, \sigma_p^2)$$

donde:

$$\mu_p = \pi \quad y \quad \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

Observación: Cuando la población tiene "N" elementos y la muestra tiene "n" elementos la desviación estándar, tanto de la media muestral como de la proporción muestral, debe multiplicarse por el factor de corrección por población.

Para la media muestral:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Para la proporción muestral:

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \sqrt{\frac{N-n}{N-1}}$$

## ACTIVIDADES

1. Las lámparas que fabrica cierta empresa tienen una vida de 800 hrs. y una desviación típica de 60 hrs. Hallar la probabilidad de que una muestra aleatoria de 16 lámparas tengan una vida media entre 769 y 810 hrs.

2. Una firma de pedidos por correo sabe por experiencias anteriores de las circulares que envía por correo, el 10% tendrán respuestas. Suponga que se envían 20 circulares como prueba de mercado en una región geográfica. Suponiendo que se puede aplicar la tasa de respuestas del 10% en la nueva región. Calcular la probabilidad que menos del 20% de la gente responde. Calcular la probabilidad que contesten entre el 25% y 35% de la gente.

3. Estudios recientemente realizados determinaron que hay una probabilidad de 0,40 que un sistema quede infectado por un virus a través de Internet. De un grupo de 50 computadoras conectadas a Internet, ¿Cuál es la probabilidad que por lo menos 23 pero no más de 25 queden infectadas.

4. Ciertos tubos fabricados por una compañía tienen una duración media de 800 horas y una desviación estándar de 60 horas. Hallar la probabilidad de que una muestra al azar de 196 tubos, tomada de entre ellos tenga una duración media de:
- 4.1 Entre 790 y 810                      4.2 menor de 785 horas

5. Se sabe que la probabilidad de falla de la tarjeta integrada de un computador es 0,12. Si se compran 50 computadoras, hallar la probabilidad que:
- 5.1 Definir la variable aleatoria a utilizar y su función de distribución de probabilidades.
- 5.2 Queden inutilizadas más de 15 computadoras.

6. La probabilidad que una persona acierte en un lanzamiento de dardos es 0,08. ¿Cuál es la probabilidad que una persona que realiza 200 lanzamientos, acierte más de 25 lanzamientos?

7. La probabilidad de ganar un premio en una máquina traga monedas es 0,05. Después de realizar 500 intentos, determinar la probabilidad de haber ganado:
- 7.1 Como mínimo 20 veces.

7.2 Más de 24 pero como máximo 30 veces.

8. Un sistema esta formado por 100 componentes que funcionan independientemente. La probabilidad que cualquier componente falle durante el periodo de operación es igual a 0,10. Si el sistema si funciona al menos 85 componentes. Calcular la probabilidad que no funcione el sistema.

9. Un jugador peruano de fútbol acierta 3 de 5 disparos penales. Suponiendo que durante toda su carrera realiza 120 lanzamientos, determine la probabilidad que anote por lo menos 45 goles.

## Autoevaluación

1. Históricamente, el 10% de un embarque grande de piezas de maquinaria es defectuosa. Suponga que el embarque consta de 5000 piezas de máquina. Si se seleccionan muestras sin reemplazo de 400 piezas. Qué proporción de las muestras tendrá:
  - a) Entre 9% y 10% de piezas defectuosas
  - b) Menos de 8% de piezas defectuosas
  - c) Por arriba de qué valor en porcentaje caerá el 95% de los porcentajes de la muestra
2. Se sabe que el 40% de las personas prefieren consumir una determinada marca de cerveza. ¿Cuál es la probabilidad que en una muestra de 144 personas, el porcentaje de personas que prefieran dicha marca de cerveza haya disminuido en más del 5%?
3. El año pasado el 10% de los dispositivos eléctricos para monitores modelo VTV - 08809 tuvieron defectos de fabricación por lo que tuvieron que ser devueltos. Si durante el presente año se han vendido 250 dispositivos, ¿Cuál es la probabilidad que se hayan devuelto menos del 8% de los dispositivos?
4. Con base a datos pasados, el 30% de las compras con tarjeta de crédito en una tienda son por cantidades superiores a \$100. Si se seleccionan muestras aleatorias de 100 compras ¿qué proporción de las muestras es posible que tengan entre el 20% y 30% de compras mayores a \$100?
6. Una firma de pedidos por correo sabe, por experiencias anteriores, que el 10% de sus envíos obtendrá respuestas. Suponga que se envían 20 circulares como prueba de mercadeo para una determinada región geográfica. Halle la probabilidad que:
  - 6.1 Menos del 20% de la gente responda
  - 6.2 Contesten entre el 20% y el 30%
7. Se sabe que el 15% de un embarque grande de piezas para maquinaria es defectuosa. Si se escogen muestras aleatorias de 400 piezas. Qué proporción de las muestras tendrá:
  - 7.1 entre 9% y 10% de piezas defectuosas.
  - 7.2 más del 8% de piezas defectuosas.
8. Se sabe que el 40% de las personas que se presentan a un empleo son altamente capacitadas en el manejo de Visual Basic. Si 25 personas se presentan a un empleo:
  - 8.1 Defina la variable aleatoria y su función de probabilidad
  - 8.2 ¿Cuál es la probabilidad que por lo menos 12 sean altamente capacitadas en el manejo de Visual Basic?
9. Determine el tamaño de muestra para que la probabilidad que un grupo de personas pese más de 75 Kg. sea igual a 1.54%. Asuma que los pesos se distribuyen normalmente con media igual a 70 Kg. y desviación estándar de 6 Kg.

10. Una máquina que empaqueta bolsas de café automáticamente, está regulada para embalar bolsas cuyos pesos se distribuyen regularmente con media 500 gr. y desviación estándar 10 gr. Para mantener la producción bajo control se selecciona una muestra aleatoria de 110 bolsas y luego se pesa. Calcular la probabilidad que el peso promedio obtenido difiera del peso promedio esperado en menos de 2 gr.
11. Las ventas mensuales de una tienda están distribuidas normalmente con media 5 000 dólares y desviación estándar de 1 000 dólares.
  - 11.1 ¿Cuál es la probabilidad que en un mes cualquiera, las ventas sean mayores a 6 500 dólares?, Interprete
  - 11.2 Si se escoge una muestra aleatoria de 6 meses (de un año determinado), ¿Cuál es la probabilidad que el promedio mensual de la muestra esté comprendido entre 4 000 dólares y 7 000 dólares? Interprete.  
Determinar la probabilidad que las ventas anuales sean mayores a 70 000 dólares. Interprete.
12. El promedio diario de ventas que realiza una tienda es 80 000 soles con una desviación estándar de 10 000 soles. Si la distribución de las ventas es normal, halla la probabilidad de:
  - 12.1 que las ventas en un día cualquiera pasen los 100 000 soles
  - 12.2 que el promedio de ventas en 100 días sea menor que 75 000 pesos.

## Para recordar.

- Es valido para cualquier población finita o infinita discreta o continua cuando  $n$  es mayor o igual a 30.
- Si la población es normal, el teorema se cumple cualquiera que sea el tamaño de la muestra.





## **CONTROL DE CALIDAD Y PROCESOS ESTADÍSTICOS**

### **TEMAS**

- Administración de la Calidad total.
- Grafica de control para medias: Gráfica X.
- Gráfica de control para la variación de procesos: Gráfica R.

### **OBJETIVOS ESPECÍFICOS**

- Reforzar los conceptos distribución Normal.

### **CONTENIDOS**

- Distribución normal.

### **ACTIVIDADES**

- Practicar con los exámenes pasados.

## ADMINISTRACION DE LA CALIDAD

Cuando pensamos en la calidad de los productos, pensamos en un conjunto de características que esperamos posea el producto. Hoy en día en los negocios se están promoviendo el concepto de **Administración de Calidad Total (ACT)**.

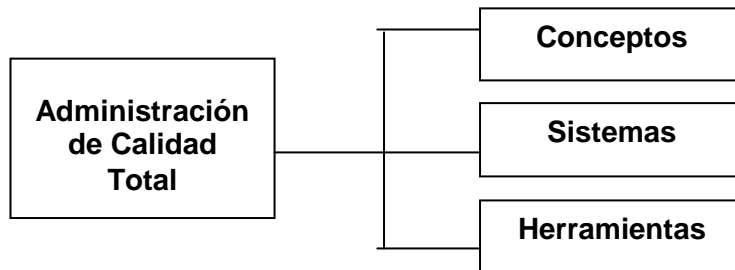
La ACT tiene tres componentes clave: (1) Conceptos, (2) Sistemas y (3) Herramientas.

El componente de conceptos de la ACT incluye varias ideas que rodean al movimiento de calidad total. Estas ideas incluyen satisfacción del cliente, todo trabajo es un proceso, hablar con datos y administración del flujo previo. Hablar con datos es un concepto de especial relevancia que implica medir y vigilar las variables de proceso.

El componente sistemas, implica la noción de administración de los sistemas. Los sistemas como gerencia general, creación de mercados, creación de productos y suministros de productos deben ser administrados en forma responsable por los encargados de la compañía.

El componente herramientas sirve para implementar un programa de ACT, incluye graficas de flujo, gráficas de causa y efectos y gráficas de control estadístico de procesos (CEP) que permite a los ingenieros entender y vigilar la variación de los procesos mediante gráficas de control.

Estos tres componentes son necesarios para implementar con éxito la filosofía de ACT en una empresa.



## Gráficas de control de variables

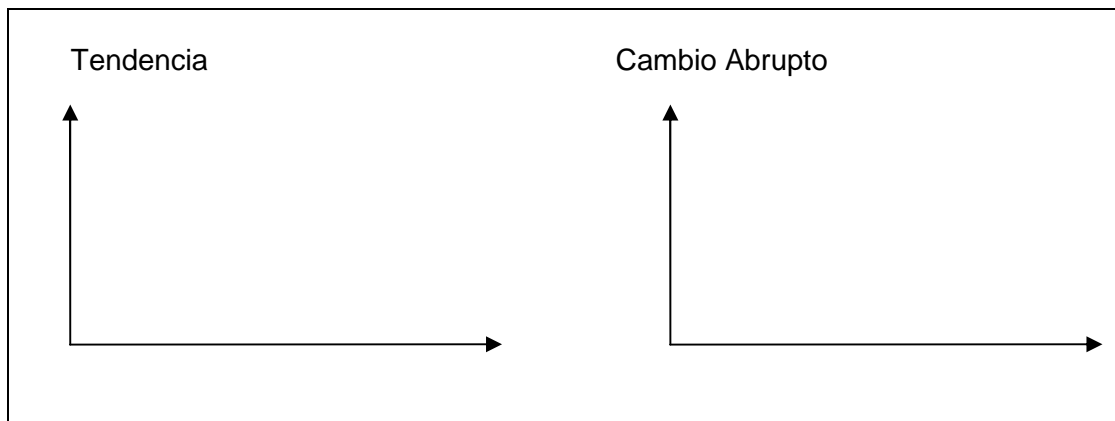
Def. Una grafica de control para una variable de calidad se obtiene graficando periódicamente las mediciones de la variable sobre el tiempo.

Las gráficas de control se construyen graficando unas variables de calidad de un producto sobre el tiempo de una grafica de secuencia, Las variable graficada puede ser una característica cuantitativa o un atributo cualitativo de un producto manufacturado. La potencia de esta grafica radica en su capacidad de separar dos tipos de variación en una característica de calidad de un producto: (1) Variación debida a causas asignables y (2) Variación aleatoria.

Las variaciones debidas a causas asignables son producidas por cosas como el desgaste de una maquina de corte metálico, el desgaste de una rueda abrasiva, cambios en la humedad y temperatura del área de producción, fatiga de los trabajadores, etc. Por lo regular se manifiestan a través de tendencias graduales en una característica con el tiempo.

Cuando la materia prima a menudo produce un cambio abrupto en el nivel de una característica de calidad. Los encargados de control de calidad y de producción intentan identificar las tendencias o los cambios abruptos en una característica de calidad cuando ocurren, y modifican el proceso a fin de reducir o eliminar este tipo de variación.

Graficas de características de calidad debida a causas asignables



**Def.** La variación de una variable que mide una característica de calidad de un producto se debe a una causa asignable o a una variación aleatoria (fortuita).

**Def.** Se dice que un proceso está en control si las características de calidad del producto están sujetas solo a variaciones aleatorias.

**Def.** La línea central de una gráfica de control variable es la media de las mediciones de la muestra, luego  $LC = \bar{x}$ .

**Def.** El límite del control interior (LCI) y el límite de control superior (LCS) para un diagrama de control variable se calculan como sigue:

$$LCI = \bar{x} - 3s$$

$$LCS = \bar{x} + 3s$$

Donde  $s$  es la desviación estándar de las mediciones de la muestra.

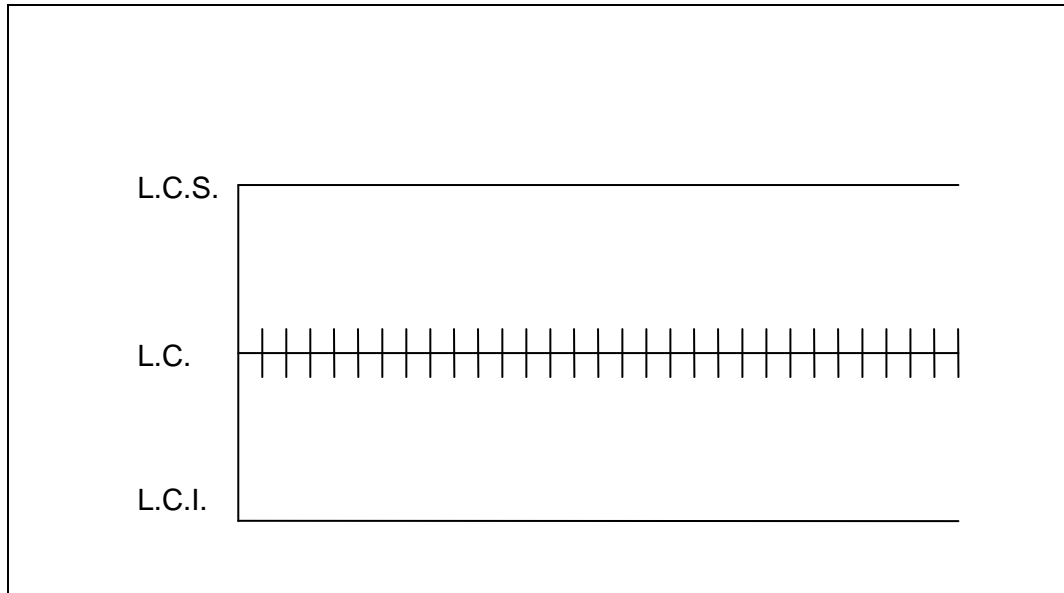
Los límites de control obtenidos de una gráfica de control sólo son apropiados para analizar datos del pasado, es decir, los datos que se emplearon para calcular los límites. Dichos límites sólo pueden aplicarse a datos futuros si el proceso está en control y/o los límites de control se modifican

## Actividades

- Una fábrica de armamento está fabricando rifles para el Ministerio de Defensa, y en una línea de producción elabora percutores terminados. A fin de vigilar el proceso, el encargado del control de calidad selecciona al azar un percutor de la línea de producción cada 30 minutos y se mide su longitud (en pulgadas). Obteniendo las siguientes observaciones.

1,00	0,99	1,01	1,06	0,99	1,01	0,99	0,99	0,98	1,03
1,01	0,99	0,98	0,99	0,86	0,99	0,97	1,01	0,97	0,87

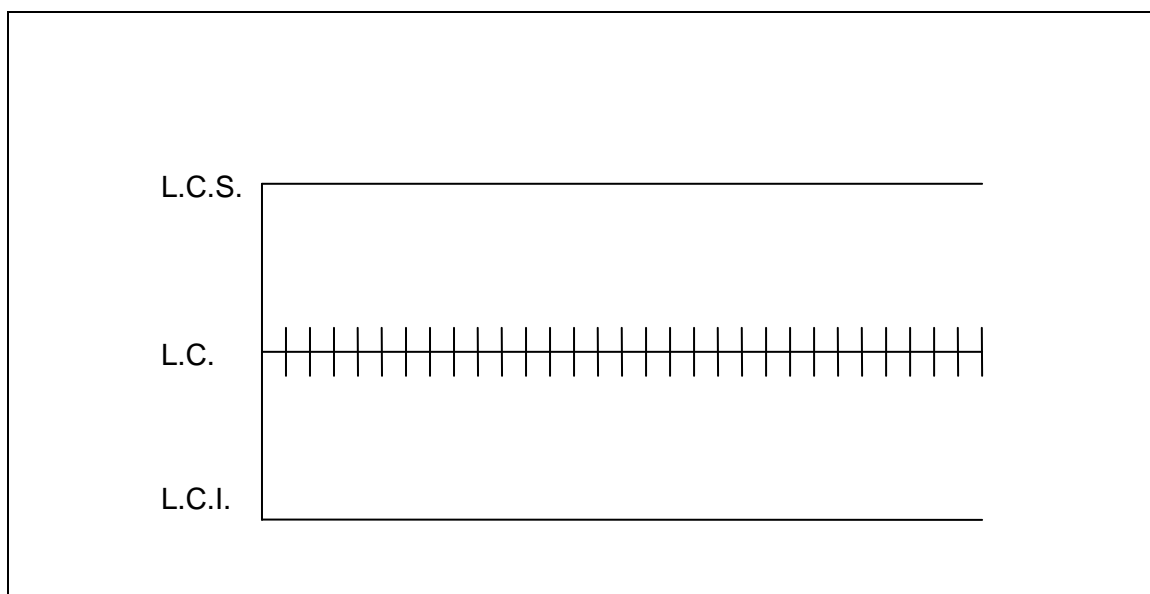
Prepare una gráfica de control para las longitudes de los percutores. ¿El proceso está fuera de control?



2. Cada día, un ingeniero de control de calidad de una compañía que fabrica botellas, selecciona al azar una botella terminada del proceso de producción en 20 puntos al mismo tiempo, además registra el peso de cada botella (en onzas). La tabla muestra los datos de dichos pesos, para la inspección del último día.

Día	1	2	3	4	5	6	7	8	9	10
<b>Peso de la Botella</b>	5,6	5,7	6,1	6,3	5,2	6,3	5,8	5,8	6,2	6,4
Día	11	12	13	14	15	16	17	18	19	20
<b>Peso de la Botella</b>	6,2	5,9	5,2	6,3	6,1	5,8	6,1	6,2	5,3	5,1

- 2.1 Trace una gráfica de control de variable para los pesos de las botellas terminadas.  
 2.2 ¿Piensa usted que el proceso haya estado en control durante el mes en cuestión?



3. Suponga que el proceso para fabricar ejes de motores eléctricos está en control. Al final de cada hora, durante un período de 24 horas, el fabricante

selecciona al azar un eje y mide su diámetro. Las mediciones (en pulgadas) para las 24 muestras se presentan en la tabla siguiente. Elabore e interprete una gráfica de control para el diámetro de los ejes.

Muestra	1	2	3	4	5	6	7	8
Diámetro (pulg.)	1,505	1,496	1,516	1,507	1,502	1,502	1,489	1,485
Muestra	9	10	11	12	13	14	15	16
Diámetro (pulg.)	1,503	1,485	1,491	1,486	1,510	1,495	1,504	1,499
Muestra	17	18	19	20	21	22	23	24
Diámetro (pulg.)	1,505	1,496	1,502	1,501	1,497	1,501	1,503	1,494

4. Ciertas juntas de expansión de hule moldeado, empleadas en sistemas de calefacción y aire acondicionado, se diseñan con un diámetro interno de 5 pulgadas. A fin de vigilar el proceso de fabricación, se seleccionó una junta al azar cada hora de la línea de producción y se midió su diámetro (en pulgadas), durante un período de 12 horas, produciendo los datos de la siguiente tabla. Elabore e interprete una gráfica de control para el diámetro de la junta de expansión de hule moldeado.

Muestra	1	2	3	4	5	6	7	8	9	10	11	12
Diámetro (pulg.)	5,08	4,88	4,99	5,04	5,00	4,83	5,02	4,91	5,06	4,92	5,01	4,92

5. La perilla de un reóstato, producida mediante moldeo de plástico, contiene un inserto metálico. El ajuste de esta perilla en su unidad depende de la distancia entre la parte posterior de la perilla y el lado lejano de un agujero cilíndrico. A fin de vigilar la operación de moldeo, se hizo un muestreo al azar de una perilla de la producción cada hora y se midió la dimensión de cada una. La siguiente tabla presenta las mediciones de distancia (en pulgadas) para las primeras 28 horas en que estuvo en operación el proceso.

Hora	Distancia	Hora	Distancia
1	0,140	15	0,144
2	0,138	16	0,140
3	0,139	17	0,137
4	0,143	18	0,137
5	0,142	19	0,142
6	0,136	20	0,136
7	0,142	21	0,142
8	0,143	22	0,139
9	0,141	23	0,140
10	0,142	24	0,134
11	0,137	25	0,138
12	0,137	26	0,140
13	0,142	27	0,145
14	0,137	28	0,136

- 5.1 Construya una gráfica de control de variable para el proceso.  
 5.2 Ubique en la gráfica la línea central, el límite de control superior y el límite de control inferior.  
 5.3 ¿Piensa usted que el proceso esté en control?

### GRÁFICA PARA CONTROL DE MEDIAS: GRÁFICA $\bar{x}$

Las gráficas de control construidas para vigilar una característica de calidad cuantitativa por lo regular se basan en muestras aleatorias de varias unidades de producto, más que en las características de unidades industriales individuales

Ubicación de la línea central y los límites de control para una gráfica  $\bar{x}$

Línea central:  $LC = \bar{\bar{x}}$       LCS:  $\bar{\bar{x}} + A_2 \bar{R}$       LCI:  $\bar{\bar{x}} - A_2 \bar{R}$

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i}{k} \quad \bar{R} = \frac{\sum_{i=1}^k R_i}{k}$$

Donde:  $K$  = Número de muestras, cada una de tamaño  $n$

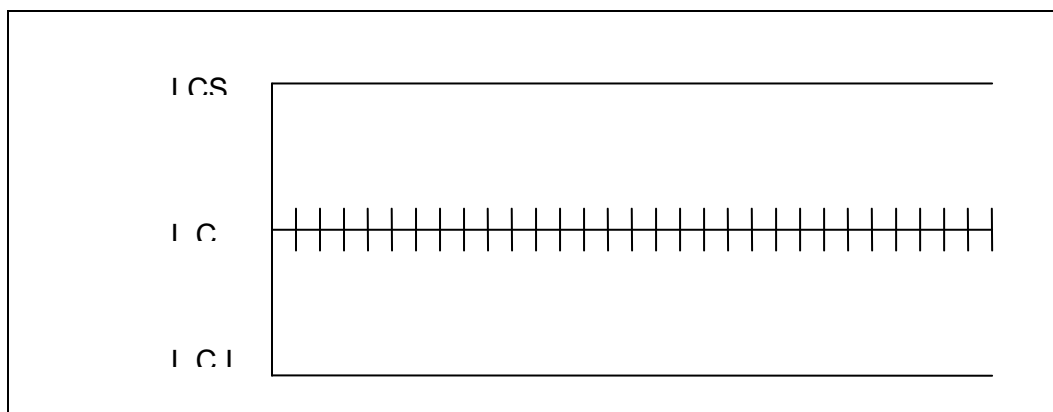
$\bar{x}_i$  = Media de la  $i$ -ésima muestra

$R_i$  = Intervalo o rango de la  $i$ -ésima muestra

Ejemplo:

Suponga que en el proceso de fabricar ejes de motores eléctricos está en control. Al final de cada hora, durante un período de 20 horas, el fabricante seleccionó una muestra aleatoria de cuatro ejes y midió el diámetro de cada uno. Las mediciones en pulgadas para 20 muestras se presentan en la siguiente tabla: Construya la gráfica de control, e interprete los resultados.

muestra	Mediciones de muestra, pulgadas				Muestra ( $\bar{x}$ )	Rango (R)
1	1,505	1,499	1,501	1,488	1,4983	0,017
2	1,496	1,513	1,512	1,501	1,5055	0,017
3	1,516	1,485	1,492	1,503	1,4990	0,031
4	1,507	1,492	1,511	1,491	1,5003	0,020
5	1,502	1,491	1,501	1,502	1,4990	0,011
6	1,502	1,488	1,506	1,483	1,4948	0,023
7	1,489	1,512	1,496	1,501	1,4995	0,023
8	1,485	1,518	1,494	1,513	1,5025	0,033
9	1,503	1,495	1,503	1,496	1,4993	0,008
10	1,485	1,519	1,503	1,507	1,5035	0,034
11	1,491	1,516	1,497	1,493	1,4993	0,025
12	1,486	1,505	1,487	1,492	1,4925	0,019
13	1,510	1,502	1,515	1,499	15065	0,016
14	1,495	1,485	1,493	1,503	1,4940	0,018
15	1,504	1,499	1,504	1,500	1,5018	0,005
16	1,499	1,503	1,508	1,497	1,5018	0,011
17	1,501	1,493	1,509	1,491	1,4985	0,018
18	1,497	1,510	1,496	1,500	1,5008	0,014
19	1,503	1,526	1,497	1,500	1,5065	0,029
20	1,494	1,501	1,508	1,519	1,5055	0,025



### Cómo interpretar una gráfica $\bar{X}$

**Proceso “Fuera de control”** Una o más de las medidas de muestra caen fuera de los límites de control. Esto indica posibles problemas en el proceso de producción y debe hacerse lo posible por determinar la causa de los valores inusualmente grandes (o pequeños) de  $\bar{X}$ .

**Proceso “en control”** Todas las medidas de muestra caen dentro de los límites de control. Aunque puede estar presente causas asignables de variación, es mejor dejar en paz el proceso y no buscar problemas que tal vez no existan.

**Def.** Los subgrupos racionales son muestras de artículos recolectados de modo que maximicen la posibilidad de que (1) las mediciones de calidad dentro de cada muestra sean similares, y (2) las mediciones de calidad entre muestras difieran.

## ACTIVIDADES:

1. Refiérase a la fabricación del proceso de ejes para motores eléctricos. Suponga que el gerente de operaciones sospecha que los trabajadores del turno nocturno están produciendo ejes con diámetros medios mayores que los trabajadores de los turnos matutino y vespertino. El gerente desea utilizar una gráfica  $\bar{X}$  para determinar si la medida del proceso ha cambiado o no. Sugiera un plan de muestreo para el gerente que siga una estrategia de subgrupos racionales.

Hora	Diámetro de paso de roscas					Hora	Diámetro de paso de roscas				
1	36	35	34	33	32	11	34	38	35	34	38
2	31	31	34	32	30	12	36	38	39	39	40
3	30	30	32	30	32	13	36	40	35	26	33
4	32	33	33	32	35	14	36	35	37	34	33
5	32	34	37	37	35	15	30	37	33	34	35
6	32	32	31	33	33	16	28	31	33	33	33
7	33	33	36	32	31	17	33	30	34	33	35
8	23	33	36	35	36	18	27	28	29	27	30
9	43	36	35	24	31	19	35	36	29	27	32
10	36	35	36	41	41	20	33	35	35	39	36

2. Refiérase a la fabricación de perillas para reóstatos. A fin de vigilar la media del proceso, se muestrearon aleatoriamente cinco perillas de cada hora de producción y se midió la distancia entre la parte de atrás de cada perilla y el fondo de un agujero cilíndrico. En la tabla se presentan las mediciones (en pulgadas) para las

primeras 24 horas de operación del proceso. Elabore e interprete una gráfica de control de medias en la fabricación de perillas de reóstatos.

Hora	Mediciones de distancia					Hora	Mediciones de distancia				
1	0,140	0,143	0,137	0,134	0,135	13	0,144	0,142	0,143	0,135	0,145
2	0,138	0,143	0,143	0,145	0,146	14	0,140	0,132	0,144	0,145	0,141
3	0,139	0,133	0,147	0,148	0,139	15	0,137	0,137	0,142	0,143	0,141
4	0,143	0,141	0,137	0,138	0,143	16	0,137	0,142	0,142	0,145	0,143
5	0,142	0,142	0,145	0,135	0,136	17	0,142	0,142	0,143	0,140	0,135
6	0,136	0,144	0,143	0,136	0,137	18	0,136	0,142	0,140	0,139	0,137
7	0,142	0,147	0,137	0,142	0,138	19	0,142	0,144	0,140	0,138	0,143
8	0,143	0,137	0,143	0,137	0,138	20	0,139	0,146	0,143	0,140	0,139
9	0,141	0,142	0,147	0,140	0,140	21	0,140	0,145	0,142	0,139	0,137
10	0,142	0,137	0,145	0,140	0,132	22	0,134	0,147	0,143	0,141	0,142
11	0,137	0,147	0,142	0,137	0,135	23	0,138	0,145	0,141	0,137	0,141
12	0,137	0,146	0,142	0,142	0,140	24	0,140	0,145	0,143	0,144	0,138

3. Una fabrica de armamento esta fabricando rifles para el Ministerio de Defensa, y en una línea de producción elabora percutores terminados. A fin de vigilar el proceso, El encargado del control de calidad selecciona al azar cinco percutores de la línea de producción cada 30 minutos y se mide su longitud (en pulgadas). Obteniendo las siguientes observaciones. Prepare una gráfica de control de medias para las longitudes de los percutores. ¿El proceso esta fuera de control?

Muestra	Mediciones de longitud (pulg.)					Muestra	Mediciones de longitud (pulg.)				
1	1,01	1,00	0,91	1,04	0,97	11	1,01	0,97	1,02	1,00	0,97
2	0,99	0,89	0,93	1,04	0,99	12	1,05	1,03	0,99	0,99	0,97
3	0,98	0,92	0,94	1,02	0,99	13	0,92	1,04	0,98	0,99	0,94
4	1,00	1,02	0,90	0,99	0,97	14	1,02	0,99	1,00	0,87	1,02
5	0,94	0,93	1,04	0,88	1,01	15	0,93	1,01	0,94	1,001	1,04
6	1,00	1,02	1,01	0,99	1,02	16	1,02	1,01	1,00	1,02	1,01
7	1,05	1,01	0,95	0,98	0,99	17	1,01	0,99	1,05	0,90	0,95
8	0,88	0,87	1,05	0,98	0,99	18	0,87	1,00	0,88	0,98	1,00
9	0,90	0,92	0,93	1,06	1,03	19	0,92	0,99	0,90	1,00	0,93
10	0,94	0,97	0,95	1,04	0,92	20	0,97	0,98	0,94	0,99	0,95

### GRAFICA DE CONTROL PARA VARIACION DE PROCESO: GRÁFICA R

En control de calidad, no solo queremos controlar el valor medio de alguna característica de calidad, sino también su variabilidad, Un incremento de la desviación estándar del proceso  $\sigma$  implica que la variable de característica de calidad varia



dentro de un intervalo más amplio, aumentando así la probabilidad de obtener un producto inferior. En consecuencia, un proceso que esta en control genera datos con una media de proceso  $\mu$  y una desviación estándar  $\sigma$  relativamente constantes.

El valor esperado y la desviación estándar del intervalo de la muestra son:

$$E(R) = d_2 \sigma \quad \sigma_R = d_3 \sigma$$

Ubicación de la línea central y los límites de control en una gráfica R

Línea central:  $LC = \bar{\bar{R}}$  donde:  $\bar{\bar{R}} = \frac{\sum_{i=1}^k R_i}{k}$

$$LCS: \bar{\bar{R}} + 3 \frac{d_3}{d_2} \bar{\bar{R}} = (1 + 3 \frac{d_3}{d_2}) \bar{\bar{R}} = D_4 \bar{\bar{R}} \quad LCI: \bar{\bar{R}} - 3 \frac{d_3}{d_2} \bar{\bar{R}} = (1 - 3 \frac{d_3}{d_2}) \bar{\bar{R}} = D_3 \bar{\bar{R}}$$

Donde:

K = Número de muestras, cada una de tamaño n

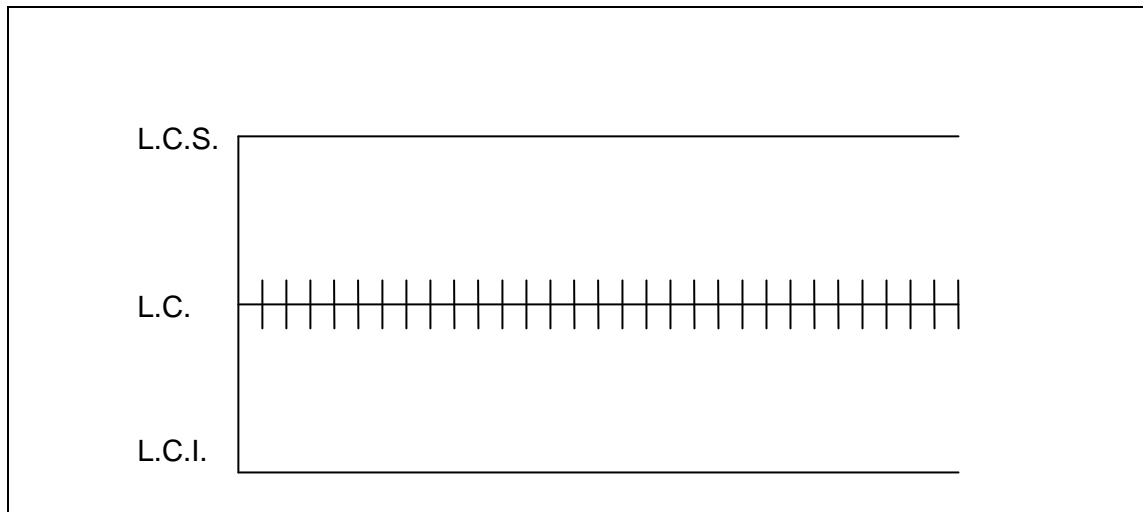
R<sub>i</sub> = Intervalo de la i-ésima muestra.

D<sub>3</sub> y D<sub>4</sub> son valores para tamaños de muestra de n = 2 hasta n = 25

Ejemplo

En el proceso de fabricación de ejes para motores eléctricos. Un fabricante seleccionó una muestra de cuatro ejes cada hora, durante un período de 20 horas, y midió el diámetro de cada uno. Suponiendo que el proceso está en control, construya e interprete una gráfica R para la variación del proceso.

Nº de muestra	Mediciones de muestra, pulgadas				Media de muestra $\bar{X}$	Rango R
1	1,504	1,499	1,502	1,482		
2	1,497	1,511	1,512	1,504		
3	1,506	1,486	1,492	1,501		
4	1,508	1,495	1,513	1,491		
5	1,506	1,491	1,501	1,503		
6	1,505	1,485	1,506	1,483		
7	1,489	1,515	1,494	1,501		
8	1,484	1,512	1,494	1,513		
9	1,503	1,494	1,503	1,494		
10	1,487	1,519	1,502	1,507		
11	1,494	1,513	1,497	1,493		
12	1,483	1,502	1,487	1,492		
13	1,512	1,504	1,515	1,496		
14	1,492	1,487	1,493	1,503		
15	1,503	1,498	1,504	1,500		
16	1,492	1,501	1,508	1,494		
17	1,503	1,497	1,509	1,491		
18	1,492	1,510	1,496	1,500		
19	1,504	1,523	1,497	1,500		
20	1,499	1,501	1,508	1,518		



### Como interpretar una gráfica R

**Proceso “Fuera de control”:** Una o más de las medidas de muestra caen fuera de los límites de control. Al igual que en el caso de la gráfica  $\bar{X}$ , esto indica un posible cambio en el proceso de producción y debe hacerse lo posible por localizar el problema.

**Proceso “en control”:** Todas las medidas de muestra caen dentro de los límites de control. En este caso, es mejor dejar en paz el proceso y no buscar problemas que talvez no existan.

## ACTIVIDADES

- Una empresa desea salir al mercado con un nuevo producto de consumo popular, a encargado a una empresa de envases de vidrio, le prepare botellas para 620 ml con un peso (en gramos) por debajo del peso de la competencia pero igualmente resistente. El fabricante hace los cambios respectivos en la cadena de producción para satisfacer los requerimientos del cliente. Para tener un control de productos terminados el fabricante ordena al encargado del control de procesos tome una muestra de 5 botellas cada hora. Construya una gráfica R, a fin de vigilar la variación en los pesos de las botellas terminadas. ¿Cree usted que el proceso esté en control?

Hora	Peso de botellas terminadas					Hora	Peso de botellas terminadas				
1	503	505	494	506	502	11	502	494	500	502	499
2	501	501	494	502	500	12	500	494	503	492	498
3	503	500	502	500	503	13	503	502	500	503	497
4	499	503	492	502	503	14	503	492	505	493	503
5	498	504	503	501	505	15	505	503	498	506	503
6	497	502	493	499	498	16	498	493	499	496	493
7	503	493	506	498	499	17	499	506	498	495	503
8	503	493	496	499	498	18	498	496	498	503	506
9	493	506	495	498	498	19	498	495	499	502	504
10	503	495	503	499	499	20	499	503	505	503	499

- Construya una gráfica R, a fin de vigilar la variación en los diámetros de paso de los aditamentos roscados. ¿Cree usted que el proceso esté en control?

Hora	Diámetro de paso de roscas					Hora	Diámetro de paso de roscas				
1	33	34	35	33	32	11	33	39	36	35	33
2	32	32	33	32	30	12	33	34	38	37	41
3	31	32	31	30	32	13	37	41	36	29	31
4	34	31	36	32	35	14	39	39	39	33	33
5	31	35	37	37	35	15	38	34	38	36	35
6	33	33	30	33	33	16	28	31	33	33	33
7	35	32	38	32	31	17	34	33	34	33	35
8	33	31	37	35	36	18	29	30	29	29	30
9	42	35	35	34	34	19	32	36	31	32	35
10	39	34	36	42	41	20	34	36	35	38	36

3. En la fabricación de perillas para reóstatos. A fin de vigilar el proceso, se muestrearon aleatoriamente cinco perillas de cada hora de producción y se midió la distancia entre la parte de atrás de cada perilla y el fondo de un agujero cilíndrico. En la tabla se presentan las mediciones (en pulgadas) para las primeras 24 horas de operación del proceso. Construya una gráfica R. ¿Cree usted que la variación del proceso esté en control?

Hora	Mediciones de distancia					Hora	Mediciones de distancia				
<b>1</b>	0,142	0,142	0,145	0,135	0,136	<b>13</b>	0,137	0,147	0,142	0,137	0,135
<b>2</b>	0,143	0,137	0,143	0,137	0,138	<b>14</b>	0,140	0,132	0,144	0,144	0,141
<b>3</b>	0,139	0,133	0,147	0,148	0,139	<b>15</b>	0,137	0,137	0,142	0,142	0,141
<b>4</b>	0,143	0,141	0,137	0,138	0,143	<b>16</b>	0,137	0,142	0,142	0,145	0,143
<b>5</b>	0,139	0,139	0,133	0,147	0,148	<b>17</b>	0,142	0,142	0,143	0,140	0,134
<b>6</b>	0,136	0,144	0,143	0,136	0,137	<b>18</b>	0,136	0,142	0,140	0,139	0,137
<b>7</b>	0,142	0,147	0,137	0,142	0,138	<b>19</b>	0,139	0,139	0,133	0,147	0,148
<b>8</b>	0,143	0,137	0,142	0,147	0,138	<b>20</b>	0,139	0,146	0,143	0,140	0,137
<b>9</b>	0,141	0,142	0,147	0,140	0,140	<b>21</b>	0,144	0,142	0,143	0,135	0,143
<b>10</b>	0,144	0,143	0,136	0,137	0,132	<b>22</b>	0,140	0,132	0,144	0,145	0,141
<b>11</b>	0,137	0,147	0,142	0,137	0,135	<b>23</b>	0,138	0,145	0,141	0,137	0,141
<b>12</b>	0,137	0,146	0,142	0,142	0,140	<b>24</b>	0,137	0,147	0,141	0,139	0,137



## **CONTROL DE CALIDAD Y PROCESOS ESTADÍSTICOS**

### **TEMAS**

- Analices de corridas.
- Gráfica de control para porcentajes de productos defectuosos: Gráfica p
- Gráfica de control para el número de defectuosos por pieza: Gráfica c
- Límites de tolerancia.

### **OBJETIVOS ESPECÍFICOS**

- Utilizar e interpretar el concepto de estimación puntual.
- Utilizar e interpretar el concepto de estimación por intervalos

### **CONTENIDOS**

- Estimación de parámetros.
- Intervalos de confianza para medias proporcionales

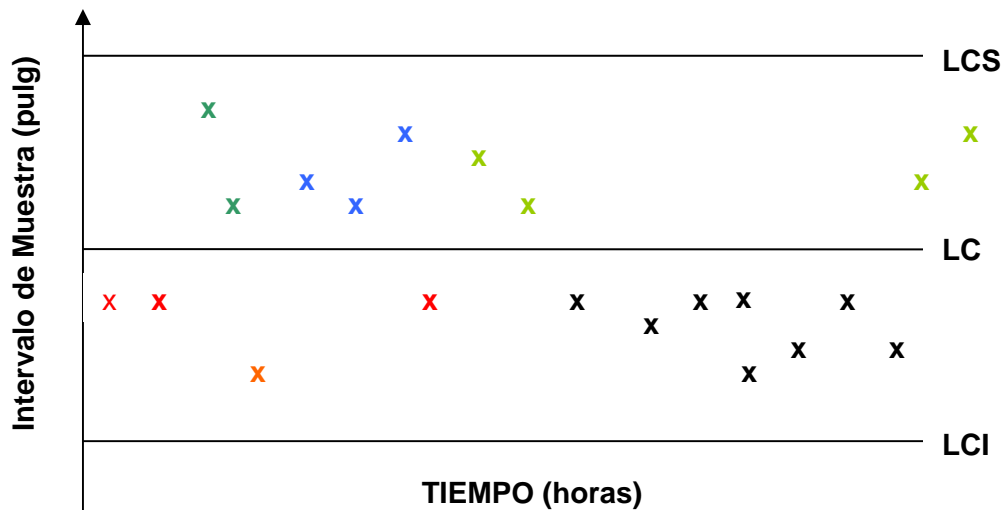
### **ACTIVIDADES**

- Estiman correcta de los parámetros

## DETECCION DE LAS TENDENCIAS EN UNA GRÁFICA DE CONTROL: ANALISIS DE CORRIDAS

Las tendencias en los procesos pueden detectarse observando corridas de puntos arriba o debajo de la línea central de la grafica de control. En control de calidad, una corrida se define como una secuencia de uno o mas puntos consecutivos, todos los cuales quedan arriba (o debajo) de la línea central.

**Def.:** Una corrida es una secuencia de uno o más puntos consecutivos que quedan del mismo lado de la línea central en una gráfica de control.



Por ejemplo se observa que la secuencia de 20 puntos consiste en un total de 8 corridas, comenzando con una corrida de 2 (-), seguida de 2 (+), 1(-), 3(+), 1(-), 2(+), 7(-) y 2(+).

En control de calidad se han desarrollado reglas para detectar tendencias que se basan únicamente en las corridas extremas (o largas) de la grafica de control.

La corrida extrema de la secuencia se compone de 7 símbolos negativos, los cuales representan los 7 intervalos de muestra consecutivos que quedan debajo de la línea central durante las 12 horas y 18 horas. La probabilidad que cualquier punto se encuentre por debajo o encima de la línea central es  $1/2$ , en forma similar la probabilidad que siete puntos consecutivos se encuentren por encima o debajo de la línea central es  $(1/2)^7$  por lo tanto la probabilidad que queden del mismo lado de la línea central es  $(1/2)^7 + (1/2)^7$  igual a 0.156

### Detección de tendencias en una gráfica de control: Análisis de corrida (o rachas)

Si se presentan cualquiera de las siguientes secuencias de corridas de grafica de control, es probable la existencia de causas de variación asignables (por ejemplo, una tendencia).

- Siete o más puntos consecutivos del mismo lado de la línea central.
- Al menos 10 de 11 puntos consecutivos en el mismo lado de la línea central.
- Al menos 12 de 14 puntos consecutivos en el mismo lado de la línea central.
- Al menos 14 de 17 puntos consecutivos en el mismo lado de la línea central.

Para los problemas propuestos de gráficas de control para medias  $\bar{X}$  y gráficas de control para variación de proceso Grafica R, realice un análisis de corrida para detectar cualquier tendencia en el proceso.

## GRAFICA DE CONTROL PARA PORCENTAJE DE PRODUCTOS DEFECTUOSOS: GRÁFICA p

Además de medir características de calidad, es importante vigilar la proporción **binomial**  $p$  de los artículos producidos que presenten defectos, luego la proporción ( $p$ ) de una muestra será igual al número de artículos defectuosos ( $y$ ) entre el número de artículos de la línea de producción ( $n$ ) al final de algún periodo específico. ( $p = y/n$ )

### UBICACIÓN DE LA LINEA CENTRAL Y LOS LÍMITES DE CONTROL DE LA GRÁFICA p

$$\bar{p} = \frac{\text{Total de artículos defectuosos}}{\text{Total de artículos inspeccionados}} = \frac{\sum_{i=1}^k p_i}{nk}$$

$$\text{LCS} : \bar{p} + 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} \quad \text{LCI} : \bar{p} - 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

Donde:

$K$  es el número de muestras, cada una de tamaño  $n$ .

$y_i$  es el número de artículos defectuosos en la  $i$ -ésima muestra.

$p_i$  es la proporción de artículos defectuosos en la  $i$ -ésima muestra.

## ACTIVIDADES:

1. A fin de vigilar el proceso de fabricación de cojinetes de apoyo de caucho empleados entre la superestructura y las bases de un cimiento de las plantas de energía nuclear, un ingeniero de control de calidad muestra aleatoriamente 100 cojinetes de la línea de producción cada día durante un periodo de 15 días. Se inspeccionaron los cojinetes y se registró el número de cojinetes defectuosos, y la proporción de defectuosos por muestra tomada. Construya una grafica p para la fracción de cojinetes defectuosos.

Día	1	2	3	4	5	6	7	8
Nº de art. defectuosos	2	12	3	4	4	1	3	5
Proporción defectuosa	0,02	0,12	0,03	0,04	0,04	0,01	0,03	0,05
Día	9	10	11	12	13	14	15	totales
Nº de art. defectuosos	3	2	10	3	3	2	3	60
Proporción defectuosa	0,03	0,02	0,10	0,03	0,03	0,02	0,03	0,04

## Cómo interpretar p

**Proceso “fuera de control”:** Una o más de las proporciones de muestra caen fuera de los límites de control. Esto indica un posible problema en el proceso de producción y amerita mayor investigación.

**Proceso “en control”:** Todas las proporciones de muestra caen dentro de los límites de control. En este caso, es mejor dejar en paz el proceso y no buscar problemas que tal vez no existan.

## ACTIVIDADES

- Los tubos cilíndricos de concreto (hormigón) preesforzado (PCCP, por sus siglas en inglés) son tubos rígidos diseñados para aprovechar al máximo la resistencia a la tensión del acero y la resistencia a la compresión y las propiedades de inhibición del concreto. El PCCP, producido en tramos de 24 pies, es propenso a un mayor agrietamiento por esfuerzos durante el proceso de fabricación. A fin de vigilar el proceso, se muestrearon 20 secciones de PCCP cada semana durante un periodo de seis semanas. En la tabla se registró el número de secciones defectuosas (esto es, el número de secciones con grietas por esfuerzo importantes).

Semana	1	2	3	4	5	6
Secciones defectuosas	1	0	2	2	3	1

- Construya una gráfica para el porcentaje de muestra de secciones de PCCP defectuosas producidas.
  - para Ubique la línea central en la gráfica p.
  - Ubique los límites de control superior e inferior en la gráfica p. ¿Cree usted que el proceso esté en control?
- Un fabricante de fusibles para terminales de computadora quiere establecer una gráfica de control para vigilar el proceso de producción. Cada hora, durante un periodo de 30 horas en el que se sabía que el proceso está en control, un ingeniero de control de calidad seleccionó al azar y probó 100 fusibles de la línea de producción. En la tabla se presenta el número de fusibles defectuosos encontrados cada hora.

Hora	1	2	3	4	5	6	7	8	9	10
Fusibles defectuosos	4	6	3	9	0	6	4	2	1	2
Hora	11	12	13	14	15	16	17	18	19	20
Fusibles defectuosos	1	3	4	5	5	2	1	1	0	3
Hora	21	22	23	24	25	26	27	28	29	30
Fusibles defectuosos	7	9	2	10	3	4	4	2	3	1



- 2.1 Construya una gráfica p para el porcentaje de muestra de fusibles para terminal defectuosos fabricados.
  - 2.2 Ubique la línea central en la gráfica.
  - 2.3 Ubique los límites de control superior e inferior en la gráfica p. ¿Cree usted que el proceso esté en control?
  - 2.4 Realice un análisis de corridas con los puntos de la gráfica p. ¿Qué implica el análisis?
3. Refiérase al ejercicio anterior. Suponga que la siguiente muestra de 100 fusibles de terminal seleccionados en línea de producción contiene 11 defectuosos. ¿El proceso está ahora fuera de control?
  4. Una compañía de productos electrónicos fabrica varios tipos de tubos de rayos catódicos mediante producción en masa. A fin de vigilar el proceso, se muestrearon aleatoriamente 50 tubos de cierto tipo de una línea de producción y se inspeccionaron. Esto se hizo diariamente durante un período de un mes. En la siguiente tabla se presenta el número de tubos defectuosos encontrados en cada día.

Día	Tubos defectuosos	Día	Tubos defectuosos
1	11	12	23
2	15	13	15
3	12	14	12
4	10	15	11
5	9	16	11
6	12	17	16
7	12	18	15
8	14	19	10
9	9	20	13
10	13	21	12
11	15	22	10

- 4.1 Construya una gráfica p para la fracción de muestra de los tubos de rayos catódicos defectuosos.
- 4.2 Ubique la línea central en la gráfica p.
- 4.3 Ubique los límites de control superior e inferior en la gráfica p.
- 4.4 ¿Cree usted que el proceso esté en control? Si no, modifique los límites de control para datos futuros.
- 4.5 Realice un análisis de corridas para detectar una tendencia en el proceso de producción.

## GRÁFICA DE CONTROL PARA EL NÚMERO DE DEFECTOS POR PIEZA: GRÁFICA C.

Otra característica de calidad es el número de defectos o imperfecciones contenidos en cada artículo individual. Por ejemplo un fabricante de muebles de oficina podría muestrear al azar un mueble de la línea de producción cada 15 minutos y registrar el número de imperfecciones en el acabado. De forma similar un fabricante de tela puede seleccionar un trozo de tela de 1m<sup>2</sup> de material seleccionado al azar cada hora y contar el número de defectos menores que contiene.

La distribución de probabilidad de **Poisson** es un buen modelo para la distribución de probabilidad del número  $c$  de defectos contenidos en un producto manufacturado.

$$E(c) = \lambda \quad \sigma_c = \sqrt{\lambda}$$

### Ubicación de la línea central y los límites de control en una gráfica c

$$\text{Línea central: } \bar{c} \quad \text{LCS: } \bar{c} + 3\sqrt{\bar{c}} \quad \text{LCI: } \bar{c} - 3\sqrt{\bar{c}}$$

Donde:

$K$  = Número de periodos de tiempo muestreados.

$C_i$  = Número de defectos por artículo observados en el tiempo  $i$ .

$$\bar{c} = \frac{\sum_{i=1}^K C_i}{K} = \text{Número medio de defectos por artículo observados durante todos los periodos de tiempo.}$$

Ejemplo.

El número de defectos perceptibles detectados por los inspectores de control de calidad en un espécimen de 1 metro cuadrado de tela de lana seleccionada aleatoriamente de cierto telar se registró cada hora durante un periodo de 20 horas. Los resultados se muestran en la siguiente tabla. Suponiendo que el número de defectos por metro cuadrado tiene aproximadamente una distribución de probabilidad de Poisson, construya una gráfica c para vigilar el proceso de producción textil.

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

**COMO INTERPRETAR UNA GRAFICA C**

Proceso “fuera de control”: Una o más de las cantidades de defectos de muestra caen fuera de los límites de control. Esto indica un posible problema en el proceso de producción y amerita mayor investigación.

Proceso “en control”: Todas las cantidades de defectos de muestra caen dentro de los límites de control. En este caso, es mejor dejar en paz el proceso y no buscar problemas que tal vez no existan.

## ACTIVIDADES

- 1 El número de imperfecciones (rayones, astilladuras, grietas y ampollas) en tableros de madera de ebanistería fabricados a la medida es importante tanto para el cliente como para el constructor. A fin de vigilar el proceso de fabricación, cada hora durante 15 horas consecutivas se selecciona un tablero acabado de 4 x 8 pies y se inspecciona para detectar imperfecciones. En la tabla se registró el número de imperfecciones del tablero.

Tablero	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Nº de defectos	4	2	3	3	9	4	5	3	8	7	3	6	5	7	5

- Grafique el número de defectos por tablero en una gráfica c.
  - Ubique la línea central de la gráfica c.
  - Ubique los límites de control superior e inferior para la gráfica c. ¿El proceso está en control?
  - Realice un análisis de corridas para la gráfica c. ¿Qué implica el análisis?
- 2 El supervisor de una operación de perforación de tarjetas para computadora realizó un estudio de control de calidad. De la producción de cada uno de los últimos 20 días se escogieron al azar que inspeccionaron 25 tarjetas perforadas para detectar errores de perforación. El número de errores observados por día se registró en la siguiente tabla.

Día	1	2	3	4	5	6	7	8	9	10
Nº de errores	7	3	9	8	2	5	10	5	7	6
Día	11	12	13	14	15	16	17	18	19	20
Nº de errores	1	4	11	8	6	6	9	2	12	9

- Grafique el número de errores de perforación por día en una gráfica c.
  - Ubique la línea central de la gráfica c.
  - Ubique los límites de control superior e inferior para la gráfica c. ¿El proceso está en control?
  - Realice un análisis de corridas para la gráfica c. ¿Qué implica el análisis?
- 3 Cierta modelo de avión es propenso a errores de alineación en el proceso de fabricación. A fin de vigilar el proceso, se registró el número total de errores de alineación observados durante la inspección final para cada uno de los primeros 26 aviones producidos, como se aprecia en la siguiente tabla.

Avión	Nº de errores de alineación	Avión	Nº de errores de alineación
1	7	14	11
2	13	15	8
3	4	16	10
4	5	17	8
5	9	18	7

<b>6</b>	3	<b>19</b>	16
<b>7</b>	4	<b>20</b>	13
<b>8</b>	6	<b>21</b>	12
<b>9</b>	7	<b>22</b>	9
<b>10</b>	14	<b>23</b>	11
<b>11</b>	18	<b>24</b>	11
<b>12</b>	11	<b>25</b>	8
<b>13</b>	11	<b>26</b>	7

- Construya la gráfica c para el número de errores de alineación por avión
  - Ubique la línea central y los límites de control superior e inferior en la grafica c.
  - ¿Cree usted que el proceso está en control? Recomienda utilizar estos límites de control para datos futuros.
4. Debido a que los datos tomados para el ensayo anterior son observados por la gerencia, estos toman el acuerdo que la prueba se realice incrementando la muestra, como se aprecia en la siguiente tabla:

Avión	Nº de errores de alineación	Avión	Nº de errores de alineación	Avión	Nº de errores de alineación	Avión	Nº de errores de alineación
<b>1</b>	7	<b>14</b>	7	<b>27</b>	9	<b>40</b>	11
<b>2</b>	6	<b>15</b>	13	<b>28</b>	8	<b>41</b>	8
<b>3</b>	6	<b>16</b>	4	<b>29</b>	15	<b>42</b>	10
<b>4</b>	7	<b>17</b>	5	<b>30</b>	6	<b>43</b>	8
<b>5</b>	4	<b>18</b>	9	<b>31</b>	4	<b>44</b>	7
<b>6</b>	7	<b>19</b>	3	<b>32</b>	13	<b>45</b>	16
<b>7</b>	8	<b>20</b>	4	<b>33</b>	7	<b>46</b>	13
<b>8</b>	12	<b>21</b>	6	<b>34</b>	8	<b>47</b>	12
<b>9</b>	9	<b>22</b>	7	<b>35</b>	15	<b>48</b>	9
<b>10</b>	9	<b>23</b>	14	<b>36</b>	6	<b>49</b>	11
<b>11</b>	8	<b>24</b>	18	<b>37</b>	6	<b>50</b>	11
<b>12</b>	5	<b>25</b>	11	<b>38</b>	10	<b>51</b>	8
<b>13</b>	5	<b>26</b>	11	<b>39</b>	9	<b>52</b>	7

- ¿Cree usted que el proceso todavía está en control?
  - Realice un análisis de corridas para la gráfica c modificada. ¿Qué detecta usted?
5. Ciertas juntas de expansión de hule moldeado, empleadas en sistemas de calefacción y aire acondicionado, se diseñan con un diámetro interno de 5 pulgadas. A fin de vigilar el proceso de fabricación, se seleccionó una junta al azar cada hora de la línea de producción y se observó un número de imperfecciones (rayones, astilladuras, grietas), durante un período de 12 horas, produciendo los datos de la siguiente tabla.
- Ubique los límites de control superior e inferior para la gráfica c. ¿El proceso está en control?

b. Realice un análisis de corridas para la gráfica c. ¿Qué implica el análisis?

Hora	1	2	3	4	5	6	7	8	9	10	11	12
N°de imperfecciones	8	2	12	14	5	7	12	15	13	17	13	12



## **REGRESIONES CORRELACIONES**

### **TEMAS**

- Análisis de Regresión Simple.
- Diagrama de dispersión. Recta de mínimos cuadrados.
- Correlación entre dos variables
- Aplicación e interpretación.

### **OBJETIVOS ESPECÍFICOS**

- Comprender los alcances de la regresión Simple y las ventajas y desventajas que presenta para realizar pronósticos. Elaborar un modelo de regresión lineal simple.
- Determinar cuantitativamente el grado de asociación entre dos variables.

### **CONTENIDOS**

- Regresión lineal simple.
- Diagrama de dispersión.
- Error estándar.
- Correlaciones

### **ACTIVIDADES**

- Interpretar la relación entre las variables independiente y dependiente; y pronosticar
- Relacionan la variable independiente y dependiente e interpretar.

# ANÁLISIS DE REGRESIÓN Y CORRELACIÓN

## Análisis de Regresión

El objetivo del análisis de regresión es el de establecer una relación cuantitativa entre dos o más variables seleccionadas. Trata de establecer un modelo que permite predecir, explicar o estimar el valor de una variable (dependiente) en función de otras variables (independientes).

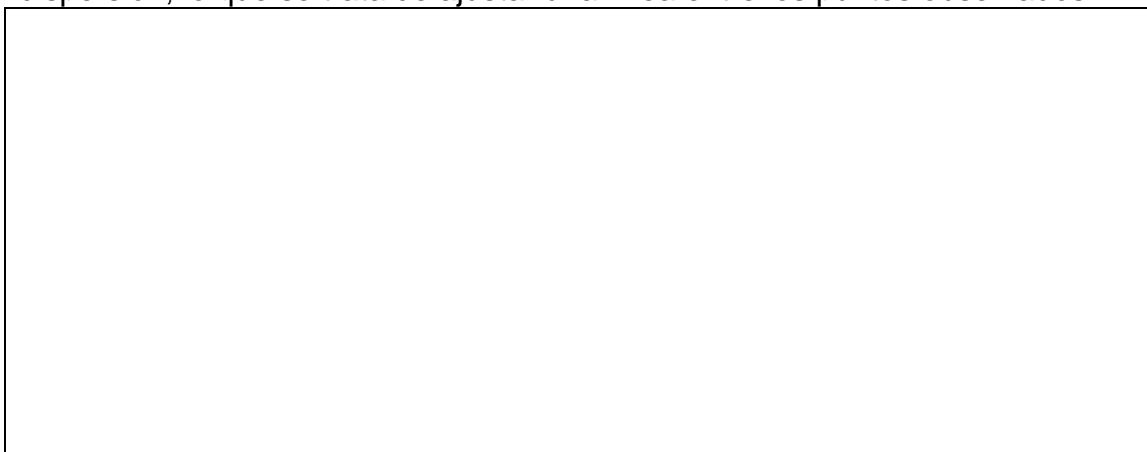
El modelo a establecer puede tener la forma de la ecuación de una recta (modelo lineal), de la ecuación de una parábola, de la función logarítmica, de la función exponencial. Asimismo, según el número de variables independientes, el modelo puede ser simple (una variable independiente) o múltiple (más de una variable independiente).

Para establecer cual es el modelo a utilizar es conveniente elaborar un Diagrama de Dispersión.



## Regresión lineal simple

Implica que debemos establecer un modelo que permita explicar a la variable Y (dependiente) en función de la variable independiente X. En un gráfico de dispersión, lo que se trata de ajustar una línea entre los puntos observados.



Supuestos del análisis de regresión lineal simple:

- a) Cada valor de las variables X e Y se distribuye normalmente
- b) Las medias de las distribuciones normales (de cada X e Y) se encuentran sobre la recta de regresión.
- c) Las desviaciones estándar correspondientes son iguales.



d) Los valores de Y son estadísticamente independientes.

El modelo de regresión lineal simple tiene la forma:  $Y = a + b.X$

Donde:

a: intercepto con el eje Y o la variación de Y que no es explicada por X

b: es la pendiente de la recta o el valor que varía Y cuando X aumenta en una unidad.

Para calcular estos coeficientes se utiliza el método de mínimos cuadrados a partir del cual se obtienen las siguientes fórmulas:

$$SC(X) = \sum x^2 - \frac{(\sum x)^2}{n} \qquad SC(Y) = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$SP(XY) = \sum xy - \frac{(\sum x)(\sum y)}{n} \qquad b = \frac{SP(XY)}{SC(X)}$$

$$a = \bar{y} - b\bar{x} \qquad r^2 = \frac{b^2 . SC(X)}{SC(Y)}$$

Al realizar la estimación de la variable dependiente Y utilizando la recta de regresión es obvio esperar que el resultado no sea exacto. Hay diferencias con respecto a los valores “reales” y a dichas diferencias se les denomina error. Por ello se define **al error estándar de estimación (Sxy)** como la desviación promedio de todas las observaciones con respecto a la recta de regresión. Se calcula de la siguiente manera:

$$S_{xy} = \sqrt{\frac{SC(y) - b^2 . SC(x)}{n - 2}}$$

## Análisis de Correlación

El análisis de correlación es un grupo de técnicas estadísticas que nos permiten medir la intensidad de la relación que puede existir entre dos variables. Nos ayuda a concluir si es que están fuertemente relacionadas o no.

El objetivo es determinar que tan intensa es la relación y utilizamos para ello el coeficiente de correlación (r). El valor de este coeficiente está comprendido entre -1 y 1. Si el valor de r tiende o está cerca de 1 diremos que la relación entre las variables en cuestión es bastante intensa y además directa; en cambio, si el valor es cercano a -1 la relación también es muy intensa pero inversa. Por otra parte, si el valor de r se acerca a cero podemos afirmar que no existe ninguna correlación.

El valor de la correlación se calcula a través de la siguiente fórmula:

$$r = \frac{.SC(x, y)}{\sqrt{SC(x) \cdot SC(y)}}$$

Una vez establecido el modelo habría que preguntarse: ¿Qué tan confiables son las predicciones que se hagan a través del modelo? Para ello es fundamental calcular el **coeficiente de determinación** ( $r^2$ ). Dicho coeficiente nos indica el porcentaje en que la variación de la variable dependiente es “explicada” por la variable independiente. Los valores de este coeficiente se encuentran comprendidos entre 0 y 1 y mientras más cercano de encuentre de 1, el modelo será más confiable. El coeficiente de determinación se calcula de la siguiente manera:

$$r^2 = \frac{b^2 \cdot SC(x)}{SC(y)}$$

## ACTIVIDADES:

1. En la siguiente tabla se muestra las ventas que tuvo una empresa en los últimos 10 años. En base a estos datos estime las ventas para el año 2001.

AÑO	VENTAS (miles de millones de dólares)
1991	0.2
1992	0.4
1993	0.5
1994	0.9
1995	1.1
1996	1.5
1997	1.3
1998	1.1
1999	1.7
2000	1.9

2. Un analista toma una muestra aleatoria a 10 embarques recientemente enviados por camión de una compañía y registra la distancia en kilómetros y el tiempo que demora la entrega desde que el embarque estuvo listo para su transportación. Estos datos se muestran a continuación:

Embarque	1	2	3	4	5	6	7	8	9	10
Tiempo (días)	3.5	1.0	4.0	2.0	1.0	3.0	4.5	1.5	3.0	5.0
Distancia(Km)	825	215	1070	550	480	920	1350	325	670	1215

- 2.1 Construya un diagrama de dispersión y determine si un análisis de regresión lineal es apropiado para relacionar la distancia en kilómetros y el tiempo que demora la entrega desde que el embarque estuvo listo para su transportación. Justifique
- 2.2 Calcule los coeficientes de correlación y determinación e interprete los resultados.
- 2.3 Mediante el método de mínimos cuadrados, halle una ecuación que permita estimar el tiempo de demora a partir de la distancia.
- 2.4 Calcule el error estándar de estimación e interprételo.

3. El tiempo en segundos en que una computadora con procesador XTF se demora en correr un sofisticado programa de regresión se cree que depende de la cantidad de KB que pueda tener el archivo “datos” necesario para que el programa determine los resultados que se esperan. Para 10 archivos distintos “datos” se obtuvo la siguiente información:

$$\begin{array}{lll} \sum x = 1450 & \sum x^2 = 218,500 & \sum y^2 = 47225 \\ \sum y = 673 & \sum xy = 101,570 & \end{array}$$

- 3.1 Determine la recta de regresión.
- 3.2 Estime el tiempo que la computadora se demorará en correr el programa cuando se utiliza un archivo “datos” de 180 KB.

4. A continuación se presenta una serie cronológica de la producción de cierto artículo (en miles) y el costo total (en miles de dólares) del mismo.

Determinar cuál de los modelos de regresión vistos en clase es el más conveniente para estimar el costo total de producción, interprete los parámetros que considere necesarios para sustentar su respuesta

Año	Producción	Costo Total
1985	12.70	8.12
1986	4.60	9.71
1987	9.60	9.77
1988	3.20	10.19
1989	15.90	6.69
1990	38.80	3.14
1991	16.90	6.13
1992	28.20	4.18
1993	11.40	7.33
1994	7.90	10.27
1995	10.60	9.66
1996	22.50	5.76
1997	37.50	3.50
1998	12.60	6.85
1999	8.60	11.29

5. El tiempo (en segundos) que demora un cajero automático en atender a un cliente se cree que se puede estimar en función del número de operaciones ( $X_1$ ) que realiza el cliente y el monto retirado ( $X_2$ , en cientos de soles). Para ello se han establecido tres modelos:

Modelo I:  $Y = 45.6 + 15.6X_1$        $SC(x) = 986.58$        $SC(Y) = 251500.56$

Modelo II:  $Y = 19.5 + 2.56X_2$        $SC(x) = 12239.84$        $SC(Y) = 251500.56$

Determine el mejor modelo, justificando adecuadamente su respuesta. Para dicho modelo estime el tiempo en segundos que el cajero demorará en atender a una persona que realiza 4 operaciones y retira en total 850 soles. (Asuma lo que considere necesario)

## Autoevaluación

1. Al estimar las ventas (Y) de un artículo en función de los precios (X) se usó una recta de mínimos cuadrados basado en una muestra de 4 datos. Si las ventas observadas fueron 10, 8, 6, 14 y si las ventas estimadas respectivas son 10.8, 8.2, 5.6, 13.4, ¿qué porcentaje de la varianza de las ventas es explicada por la recta de regresión?.
2. En un estudio de la relación entre ingresos mensuales y gastos de educación de las familias, una muestra proporciona un coeficiente de determinación del 90.25%, medias respectivas de \$ 420 y \$ 120, y desviaciones estándar respectivas de \$10 y \$ 7. Según este estudio.
  - 2.1 ¿En cuanto se estima los gastos por educación de una familia cuyo ingreso mensual es de \$300?
  - 2.2 Si una familia estima su gasto por educación en \$ 370, ¿cuánto debería ser su ingreso mensual?
3. Un estudio de Mercado trata de averiguar si es efectiva la propaganda televisada de un producto que salió a la venta con relación al tiempo de publicidad (en horas / semana). Se recopilaron datos a partir de la segunda semana de iniciada la publicidad resultando el cuadro que sigue. No se pudo recopilar datos de la cuarta parte.

Semana	2	3	4	5	6	7
Tiempo de propaganda	20	25	22	28	36	40
Venta del producto (\$)	300	310	-	320	350	420

- 3.1 ¿Halle la ecuación de regresión lineal?.
- 3.2 ¿En cuanto estimaría las ventas para la semana 4?
- 4 Se quiere estudiar la relación entre las edades en años (X) de un tipo de máquinas que se utiliza en la fabricación de cierto artículo y el número de artículos (Y) que producen. A partir de la muestra siguiente:

X	Y
2	95
3	70, 80
4	-
5	75
6	60
7	-
8	-
9	45, 50
10	25

- 4.1 Determinar la recta de regresión de mínimos cuadrados para predecir la producción. Estimar la producción para 4, 7 y 8 años.

- 4.2 Calcular el porcentaje de la varianza explicada por la regresión de la producción.
- 4.3 Si realmente cada máquina de la muestra produce 10 artículos menos determinar la recta de regresión. ¿Cuánto es el porcentaje de la varianza explicada por la regresión de la producción?
- 5 Sea  $Y$  el índice de precios al consumidor, tomando como base el año 1990 (es decir 1990 = 100). Para los datos que siguen:

Año	1991	1992	1993	1994	1995	1996	1997
$Y$	106.0	111.1	117.2	121.3	125.2	128.0	132.6

- 5.1 Hallar la recta de mínimos cuadrados que se ajuste a los datos.
- 5.2 Predecir el índice de precios para el año 1998 y compararlo con el valor verdadero (144.4). ¿En que año podemos esperar que el índice de precios sea 150.57, suponiendo que las tendencias presentes continúen?
6. Un editor tomó una muestra de 7 libros anotando el precio y el número de páginas con el fin de predecir precios. ¿Qué porcentaje de la varianza total de precios se explica por esta función?

Precio	12	14.	19	25	30
# Pág.	175	230	340	445	560

- 6.1 Determinar el coeficiente de correlación entre el precio y el número de páginas.
- 6.2 Estimar el precio de un libro de 600 páginas. Si a este libro se le incrementa 20 paginas en una segunda edición, ¿en cuánto se incrementará su precio?
- 6.3 Cuántas páginas debería tener un libro cuyo precio se estima en \$ 22.27.
7. Una muestra de 5 varones adultos de quienes se observaron las estaturas ( $X$  en pies, pulgadas y los pesos ( $Y$  en libras) ha dado los siguientes resultados:

$X$	5'1"	5'2"	5'3"	5'4"	5'5"
$Y$	125	130	140	145	160

- 7.1 Realice una regresión lineal y utilice los datos para verificar que la varianza total de  $Y$  es igual a la varianza residual más la varianza explicada por la recta de regresión.
- 7.2 Usando la descomposición de la varianza calcule  $r^2$  e interprete el resultado.
8. Un director de personal de una Consultora en Sistemas realizó una prueba de conocimientos en programación a cinco empleados que ya laboraban y pretende

establecer en el futuro el sueldo inicial de los nuevos contratados en función al resultado obtenido en dicha prueba. Los resultados obtenidos son los siguientes:

Empleado	Puntaje	Sueldo (\$)
A	4	5000
B	7	12000
C	3	4000
D	6	8000
E	10	11000

Establecer la ecuación de regresión respectiva. (Sugerencia: el exprese el sueldo en miles de dólares).

Determine el coeficiente de determinación y de correlación, interpréte los.

## Para recordar.

- Se debe tener cuidado en definir la variable independiente y la variable dependiente para tener una correcta regresión.
- No necesariamente los puntos de dispersión pueden asegurar la trayectoria de la ecuación de regresión lineal