

INDICE GENERAL	Página
<i>PRESENTACIÓN</i>	8
I.- CONCEPTOS BÁSICOS	9
1.1.- Algunos Conceptos útiles.....	9
1.1.1.- Razón.....	9
1.1.2.- Proporción.....	9
1.1.3.- Porcentaje.....	9
1.1.4.- Tasa.....	9
1.1.5.- El Operador Sumatoria.....	9
1.1.5.1.- Número de términos de una sumatoria.....	9
1.1.5.2.- Propiedades operativas de las sumatorias.....	10
1.1.5.2.1.- Primera propiedad.....	11
1.1.5.2.2.- Segunda propiedad	11
1.1.5.2.3.- Tercera propiedad.....	11
1.1.5.2.4.- Cuarta propiedad.....	11
1.1.5.2.5.- Quinta propiedad.....	12
1.1.5.2.6.- Sexta propiedad.....	12
1.2.- Definición de Estadística.....	12
1.2.1.- Descripción de un gran número de datos.....	12
1.2.2.- Análisis de la población a través de muestras.....	12
1.2.3.- Predicción o comportamiento de los fenómenos en el futuro.....	13
1.3.- Terminología Estadística.....	13
1.3.1.- Población.....	13
1.3.2.- Elementos.....	13
1.3.3.- Caracteres.....	13
1.3.3.1.- Variables.....	13
1.3.3.2.- Atributos.....	13
1.3.4.- Censo.....	13
1.3.5.- Sub-Población.....	13
1.3.6.- Muestra.....	14
1.3.6.1.- Muestreo aleatorio.....	14
1.3.6.1.1.- Muestreo aleatorio simple.....	14
1.3.6.1.2.- Muestreo aleatorio estratificado.....	14
1.3.6.1.2.1.- Proporcional.....	15
1.3.6.1.2.2.- No proporcional.....	16
1.3.7.- Registros o Estadísticas.....	16
1.3.8.- Encuestas.....	17
1.3.9.- Estadístico.....	17
1.3.10.- Parámetro.....	17
1.3.11.- Razonamiento deductivo.....	17
1.3.12.- Razonamiento inductivo.....	17
1.4.- Pasos a seguir en una Investigación Estadística.....	17
1.4.1.- La recolección de los datos.....	17
1.4.1.1.- Preparación del cuestionario.....	19
1.4.2.- La organización de los datos.....	21
1.4.2.1.- Corrección.....	21
1.4.2.2.- Clasificación.....	21
1.4.3.- La presentación de los datos.....	21

1.4.3.1.- Mediante enunciados.....	21
1.4.3.2.- Mediante tablas estadísticas.....	21
1.4.3.3.- Mediante gráficas estadísticas.....	21
1.4.4.- El análisis de los datos.....	22
1.4.5.- La interpretación de los resultados.....	22
1.5.- El método científico.....	22
1.5.1.- Características de un Experimento bien planeado.....	22
1.5.2.- Procedimiento para la Experimentación.....	22
Ejercicios propuestos.....	24
II LA ELABORACIÓN DE TABLAS O CUADROS ESTADÍSTICOS.....	27
2.1.- Partes de un cuadro estadístico.....	27
A.- Principales.....	28
B.- Secundarias.....	28
2.2.- Tipos de cuadros estadísticos más usuales.....	29
2.2.1.- De variables.....	29
2.2.1.1.- De una sola variable.....	29
a) Que se trate de pocos elementos.....	29
b) Que se trate de muchos elementos, pero pocos valores distintos de la variable.....	30
c) Que se trate de muchos elementos y muchos valores distintos de la variable.....	30
c.1.- Cálculo de los centros de clases.....	33
c.2.- Cálculo de las frecuencias relativas.....	33
c.3.- Cálculo de las frecuencias absolutas acumuladas.	34
c.4.- Cálculo de las frecuencias relativas acumuladas.....	34
2.2.1.2.- De dos variables.....	35
a) Que se trate de pocos elementos.....	35
b) Que se trate de muchos elementos, pero pocos pares de valores distintos de las variables	35
c) Que se trate de muchos elementos y muchos pares de valores distintos de las dos variables.....	36
2.2.1.3.- De tres o más variables.....	37
a) Si se trata de pocos elementos.....	37
b) Si se trata de muchos elementos.....	38
2.2.2.- De atributos.....	38
2.2.2.1.- De un solo atributo.....	38
2.2.2.2.- De dos atributos.....	39
2.2.2.3.- De más de dos atributos.....	39
2.2.3.- Mixtas.....	39
2.2.4.- Geográficas.....	40
2.2.5.- Sectoriales.....	40
2.2.6.- Temporales.....	41
Ejercicios propuestos.....	42
III .LA ELABORACIÓN DE GRÁFICOS ESTADÍSTICOS.....	45
3.1.- Partes de un gráfico estadístico.....	45
3.2.- Elaboración de gráficos en escala aritmética.....	46
3.2.1.- El histograma de frecuencias.....	46
3.2.1.1.- Que la amplitud de las clases sea constante.....	46
3.2.1.2.- Que la amplitud de las clases no sea constante.....	47
3.2.2.- El polígono de frecuencias.....	48

3.2.2.1.- Cuando la amplitud de clases es constante.....	48
3.2.2.2.- Cuando la amplitud de clases no es constante.....	49
3.2.3.- Gráficas de partes componentes.....	50
3.2.3.1.- Simples.....	51
3.2.3.1.1.-En líneas.....	51
3.2.3.1.2.-En barras.....	52
3.2.3.2.- En porcentaje.....	53
3.2.3.2.1.-En líneas.....	53
3.2.3.2.2.-En barras.....	54
3.2.4.- Gráficas de barras.....	55
3.2.4.1.- Gráficas de barras simples.....	55
3.2.4.2.- Gráficas de barras dobles.....	56
3.2.5.- Gráficas de torta sectoriales o de pastel.....	56
3.2.6.- Gráficas de líneas.....	57
Ejercicios propuestos.....	60
IV.- MEDIDAS DE TENDENCIA CENTRAL.....	63
4.1.- La media aritmética.....	63
4.1.1.- Simple.....	63
4.1.1.1.- Para datos no repetidos ni agrupados.....	63
4.1.1.2.- Para datos repetidos.....	64
4.1.1.3.- Para datos agrupados.....	64
4.1.2.- Ponderada.....	65
4.1.3.- Propiedades de la media aritmética.....	65
4.2.- La mediana.....	67
4.2.1.- Para datos no repetidos ni agrupados.	67
4.2.1.1.- Número impar de datos.....	67
4.2.1.2.- Número par de datos.....	68
4.2.2.- Para datos repetidos.....	68
4.2.2.1.- Que $N/2$ o $(N+1)/2$ no coincida con alguna F_j	68
4.2.2.2.- Que $N/2$ o $(N+1)/2$ coincida con alguna F	69
4.2.3.- Para datos agrupados.....	70
4.2.4.- Ventajas de la mediana.....	71
4.3.- La moda.....	71
4.3.1.- Para datos no repetidos ni agrupados.....	72
4.3.2.- Para datos repetidos.....	72
4.3.3.- Para datos agrupados.....	72
4.4.- La media geométrica.....	73
4.4.1.-Para datos no repetidos ni agrupados	73
4.4.2.-Para datos Repetidos o Agrupados.....	74
4.5.- Formas que puede tomar una distribución de frecuencias.....	74
4.5.1.- Relación empírica entre media, mediana y moda.....	75
4.5.2.- Cálculo del tipo y grado de asimetría en una distribución de frecuencias campaniforme.....	75
4.6.- Cuartiles, deciles y percentiles.....	76
4.6.1.- Cuartiles.....	76
4.6.2.- Deciles.....	77
4.6.3.- Percentiles.....	78
Ejercicios propuestos.....	79

V.- MEDIDAS DE DISPERSIÓN.....	83
5.1.- La oscilación, rango o recorrido.....	83
5.2.- La desviación media absoluta.....	84
5.2.1.- Para datos no repetidos ni agrupados.....	84
5.2.2.- Para datos repetidos o agrupados.....	84
5.3.- La varianza.....	86
5.3.1.- Para datos no repetidos ni agrupados.....	86
5.3.2.- Para datos repetidos o agrupados.....	87
5.3.3.- Propiedades de la varianza.....	88
5.4.- La desviación típica o estándar.....	89
5.5.- El coeficiente de variación.....	89
5.6.- La desviación cuartílica.....	91
5.7.- Medidas de apuntamiento o kurtosis.....	91
Ejercicios propuestos.....	93
 VI.- INTRODUCCIÓN AL CÁLCULO DE LAS	
PROBABILIDADES.....	95
6.1.- La teoría combinatoria.....	95
6.1.1.- Las variaciones.....	95
6.1.2.- Las combinaciones.....	96
6.1.3.- Las permutaciones.....	96
6.2.- Noción de probabilidad.....	96
6.3.- Espacio muestral y sucesos.....	97
6.3.1.- Espacio muestral.....	97
6.3.2.- Suceso o evento.....	98
6.3.3.- Operaciones con sucesos.....	98
6.3.4.- Sucesos mutuamente excluyentes.....	99
6.3.5.- Sucesos independientes.....	99
6.4.- Axiomas de la probabilidad.....	100
6.5.- Probabilidad condicional.....	101
6.6.- Distribución de probabilidad.....	101
6.6.1.- Distribución de probabilidad discreta.....	102
6.6.2.- Distribución de probabilidad continua.....	104
6.6.3.- La distribución normal.....	105
6.6.3.1.- Generalidades.....	105
6.6.3.2.- Curva normal estandarizada.....	105
6.6.3.3.- Uso de la tabla de áreas de la curva normal estandarizada.....	107
6.6.4.- La distribución binomial.....	111
6.6.4.1.- Generalidades.....	111
6.6.4.2.- Cálculo de la media y la desviación típica en una distribución binomial.....	112
6.6.5.- La distribución de Poisson.....	112
6.6.5.1.- Generalidades.....	112
6.6.5.2.- La media y la desviación típica en la distribución de Poisson.....	113
Ejercicios propuestos.....	114
 VII.- ESTIMACIÓN DE PARÁMETROS.....	117
7.1.- Distribución en el muestreo.....	117
7.1.1.- Distribución en el muestreo de la media.....	117
7.1.1.1.- Teorema central del límite.....	117

7.1.1.2.- Ley de los grandes números.....	118
7.1.2.- Distribución en el muestreo de la proporción.....	119
7.1.3.- Distribución en el muestreo de la diferencia entre dos medias muestrales	120
7.1.4.- Distribución en el muestreo de la diferencia entre dos proporciones muestrales.....	121
7.2.- Teoría de la estimación.....	122
7.2.1.- Propiedades de un buen estimador.....	122
7.2.2.- Estimación puntual.....	123
7.2.3.- Estimación por intervalos.....	123
7.2.3.1.- Estimación de una media poblacional a partir de una media muestral ($\bar{X}$).....	123
7.2.3.2.- Estimación de una proporción poblacional a partir de una proporción muestral (p).....	125
7.3.- Determinación del tamaño apropiado de una muestra.....	126
7.3.1.- Tamaño de muestra para estimar una media poblacional.....	126
7.3.2.- Tamaño de muestra para estimar una proporción poblacional.....	127
Ejercicios propuestos.....	128
VIII.- DOCIMASIA, PRUEBA O CONTRASTE DE HIPÓTESIS.....	131
8.1.- Clases de hipótesis.....	131
8.1.1.- De acuerdo a su alcance.....	131
8.1.2.- De acuerdo a su afirmación.....	131
8.2.- Dósimas de una y dos colas. Errores en la decisión. Punto crítico.....	132
8.2.1.- Errores en la decisión.....	133
8.2.1.1.- Error tipo I (α) y tipo II (β).....	133
8.2.1.2.- Nivel de significación y punto crítico.....	133
8.2.2.- Ilustración para dósimas de una y dos colas.....	133
8.2.2.1.- Para dósimas de una cola.....	133
8.2.2.2.- Para dósimas de dos colas.....	134
8.3.- Contraste de hipótesis para la diferencia entre una media muestral y una media poblacional para muestras grandes.....	134
8.4.- Contraste de hipótesis para la diferencia entre una media muestral y una media poblacional para muestras pequeñas.....	135
8.5.- Contraste de hipótesis para la diferencia entre dos medias muestrales.....	136
8.6.- Contraste de hipótesis para la diferencia entre una proporción muestral y una proporción poblacional.....	137
8.7.- Contraste de hipótesis para la diferencia entre dos proporciones muestrales.....	137
Ejercicios propuestos.....	139
IX.- ANÁLISIS DE CORRELACIÓN Y ALGUNOS MODELOS DE REGRESIÓN.....	142
9.1.- Algunos modelos de regresión.....	142
9.1.1.- Modelo de regresión lineal.....	143
9.1.2.- Modelo de regresión exponencial.....	143
9.1.3.- Modelo de regresión geométrico o potencial.....	143
9.2.- Tipo y magnitud de la relación entre dos variables. Cálculo del coeficiente de correlación.....	143
9.2.1.- Cálculo de p para el modelo de regresión lineal.....	143

9.2.2.- Cálculo de p para el modelo de regresión exponencial.....	145
9.2.3.- Cálculo de p para el modelo de regresión potencial.....	147
9.2.4.- Propiedades del coeficiente de correlación p	148
9.3.- Cálculo de los parámetros en los modelos de regresión.....	150
9.3.1.- En el modelo de regresión lineal.....	150
9.3.2.- En el modelo de regresión exponencial.....	151
9.3.3.- En el modelo de regresión potencial.....	153
9.4.- Cálculo del error estándar de estimación.....	154
Ejercicios propuestos.....	158
 X.- INTRODUCCIÓN AL DISEÑO EXPERIMENTAL Y AL	
ANÁLISIS DE LA VARIANCIA.....	161
10.1.- El diseño experimental.....	161
10.2.- Procedimiento para el diseño de experimentos.....	162
10.3.- Experimentos con un solo factor. Análisis de la variancia.....	163
10.3.1.- Análisis de la variancia (varianza) para el diseño experimental de efectos fijos totalmente aleatorizado.....	164
10.3.2.- Análisis de la variancia (varianza) para el diseño experimental de bloques completos al azar.....	169
Ejercicios propuestos.....	174
 XI.-PRUEBAS NO PARAMÉTRICAS.....	
11.1.- Escalas de medición.....	178
11.1.1.- Escala nominal.....	178
11.1.2.- Escala ordinal.....	179
11.1.3.- Escala de intervalo.....	179
11.1.4.- Escala de proporción.....	179
11.2.- Pruebas no paramétricas para el caso de una muestra.....	179
11.2.1.- La prueba J_i cuadrado.....	179
11.2.2.- La prueba de Kolmogorov- Smirnov para una muestra.....	181
11.2.3.- La prueba de Rachas.....	183
11.3.- Pruebas no paramétricas para el caso de dos muestras relacionadas.....	185
11.3.1.- La prueba de McNemar.....	185
11.3.2.- La prueba de los signos.....	186
11.3.3.- La prueba de Wilcoxon.....	188
11.4.- Pruebas no paramétricas para el caso de dos muestras independientes.....	189
11.4.1.- La prueba de Fisher.....	189
11.4.2.- La prueba de J_i cuadrado para dos muestras independientes.....	190
11.4.3.- La prueba de la Mediana.....	192
11.4.4.- La prueba U de Mann-Whitney.....	194
11.4.5.- La prueba de Kolmogorov- Smirnov para dos muestras independientes.....	195
11.5.- Pruebas no paramétricas para el caso de k muestras relacionadas.....	196
11.5.1.- La prueba de Cochran.....	196
11.5.2.- La prueba de Friedman.....	198
11.6.- Pruebas no paramétricas para el caso de k muestras independientes.....	199
11.6.1.- La prueba de J_i cuadrado para k muestras independientes.....	199
11.6.2.- La prueba de Kruskal-Wallis.....	200
11.7.- Correlación por rangos de Spearman.....	201
Ejercicios propuestos.....	204

XII.- NÚMEROS ÍNDICES.....	211
12.1.- Clases de números índices.....	211
12.1.1.- Índices de precios.....	211
12.1.2.- Índices de cantidad.....	211
12.1.3.-Índices de productividad.....	211
12.2.- Cálculo de números índices.....	211
12.2.1.- Números índices simples.....	211
12.2.1.1.- Para un solo artículo.....	211
12.2.1.2.- Para varios artículos.....	213
12.2.2.- Números índices ponderados.....	214
12.2.2.1.- Números índices ponderados de precios.....	214
12.2.2.1.1.- Índice de precios de Laspeyres.....	214
12.2.2.1.2.- Índice de precios de Paasche.....	215
12.2.2.1.3.- Índice ideal de Fischer.....	216
12.2.2.2.- Números índices ponderados de cantidades.....	216
12.2.2.2.1.- Índice de cantidad de Laspeyres.....	216
12.2.2.2.2.- Índice de cantidad de Paasche.....	216
12.2.2.2.3.- Índice de cantidad de Fischer.....	216
12.2.2.3.- Números índices ponderados de valor.....	217
12.2.2.4.- Números índices ponderados de productividad.....	218
12.2.2.5.- Índice de poder adquisitivo.....	219
Ejercicios propuestos.....	220
 APÉNDICES.....	 222
Apéndice 1.- Áreas bajo la curva normal tipificada.....	222
Apéndice 2.- Valores de $t_{\alpha, v}$ (t de Student).....	224
Apéndice 3.- Distribución F.....	225
Apéndice 4.- Valores de $X^2_{\alpha, v}$ (Ji cuadrado).....	228
Apéndice 5.- Valores críticos de D en la prueba de una muestra de Kolmogorov-Smirnov.....	229
Apéndice 6.- Valores críticos de r en la prueba de rachas.....	230
Apéndice 7.- Valores críticos de T en la prueba de Wilcoxon.....	232
Apéndice 8.- Valores críticos de D en la prueba de Fisher.....	233
Apéndice 9.- Valores críticos de U en la prueba de Mann-Whitney.....	251
Apéndice 10.-Valores críticos de D en la prueba de Kolmogorov-Smirnov.	256
Apéndice 11.-Valores críticos de X^2_r de la prueba de Friedman.....	259
Apéndice 12.-Valores críticos de H de la prueba de Kruskal-Wallis para $k = 3$	260
Apéndice 13.-Distribución binomial.....	262
 BIBLIOGRAFÍA.....	 265

PRESENTACIÓN

Esta obra recoge el fruto de la experiencia y la consulta a lo largo de veintidós años de enseñanza del curso de Estadística a los estudiantes de las carreras de Economía Agrícola, Sociología del Desarrollo, Educación Integral y Contaduría pública de la Universidad Nacional Experimental de los Llanos “Ezequiel Zamora”, UNELLEZ, de Barinas, Venezuela, del curso de Estadística General a los estudiantes de Ingeniería Agroindustrial y Tecnología de los Alimentos y del curso de Estadística Aplicada a los estudiantes de maestría en Docencia Universitaria de la misma universidad en su vicerrectorado de la provincia de Cojedes. Se trata de una versión que contiene ejemplos y ejercicios lo más cercanos al ambiente en el cual les tocará o les toca actuar a estos estudiantes en su ejercicio profesional.

Gran parte de los ejemplos mostrados y de los ejercicios propuestos han sido desarrollados por el autor en el aula de clases.

El lenguaje utilizado es simple y sin rebuscamientos. Las operaciones matemáticas, carentes de complicación. Por ello esta obra puede ser utilizada como libro de texto o de autoaprendizaje tanto por estudiantes universitarios, como por cualquier persona que domine los elementos básicos del cálculo matemático.

Se pretende adentrar al lector en el conocimiento de una serie de herramientas y técnicas que le permitan recoger información de los procesos para poder describirlos y compararlos.

¿Cómo recabar datos?, ¿Cómo reducirlos?, ¿Cómo presentarlos?, ¿Cómo analizarlos?, pero también: ¿Cómo encontrar la relación entre dos grupos distintos de datos, si la hubiere?, ¿Cómo saber si dos grupos, aparentemente distintos de datos, en realidad no lo son?, ¿Cómo conocer el valor más probable de una medida en una población cuando solo se tiene el valor en una muestra, por lo general pequeña?, son algunas preguntas que encontrarán respuesta en estas páginas.

Nuestra intención es básicamente, dotar a los estudiantes universitarios de los elementos para su inicio en el estudio de la Estadística, elementos éstos, que les permitan encontrar un camino expedito para ordenar y procesar la información, planificar algunos experimentos y adentrarse en los cánones del control de la calidad y del control de los procesos, que, con seguridad, les tocará dirigir.

Ojalá pueda cumplir con este ambicioso objetivo.

Valencia, mayo de 2024

Héctor Rodríguez Márquez

I.- CONCEPTOS BÁSICOS

1.1.- ALGUNOS CONCEPTOS ÚTILES

1.1.5.-El operador sumatoria: En el cálculo matemático existen innumerables *operadores*; vale decir, símbolos que significan determinadas operaciones a realizar por parte de la persona ejecutante con las cifras que le son suministradas como insumo al operador. Entre otros operadores matemáticos ya famosos están: El operador suma (+), el operador resta (-), el operador multiplicación (x), el operador división (/), el operador derivada (∂), el operador integral (∫) y el *operador sumatoria* ($\sum$).

Supongamos que la variable “X” toma los valores X1, X2, X3, X4, X5, X6, luego la suma se puede expresar como:

$$X + X + X + X + X + X = \sum_{i=1}^6 X_i$$

donde “ $\sum_{i=1}^6$ ” es la letra griega sigma mayúscula y se lee: “sumatoria de X_i desde $i = 1$ hasta 6. Es decir que “ $\sum$ ” expresa una suma.

Cuando se tiene una suma de n términos tales como:

$$X + X + X + X + \dots + X_n, \text{ entonces } \sum_{i=1}^n X_i = X + X + X + X + \dots + X$$

donde “i” es el índice de la sumatoria (Puede emplearse cualquier otra letra) y tiene como principal característica la de variar adoptando sólo valores enteros consecutivos y comprendidos entre el que se indica abajo (límite inferior) y el que se indica arriba (límite superior) del signo “ $\sum$ ” de sumatoria.

Ejemplo 1.5: Si $X = 2$; $X = 7$; $X = 6$; $X = 4$; $X = 12$

$$\sum_{i=1}^5 X_i = X + X + X + X + X = 2 + 7 + 6 + 4 + 12 = 31$$

aquí “i” varía desde 1 hasta 5. También podemos calcular, por ejemplo:

$$a) \sum_{i=1}^3 X_i = X + X + X = 2 + 7 + 6 = 15$$

$$b) \sum_{i=3}^5 X_i = X + X + X = 6 + 4 + 12 = 22$$

$$c) \sum_{i=2}^5 X_i = X + X + X + X = 7 + 6 + 4 + 12 = 29$$

Otros ejemplos son:

$$d) 2 + 4 + 6 + 8 = \sum_{i=1}^4 2i$$

$$e) 2 + 4 + 8 + 16 + 32 = \sum_{a=1}^5 2^a$$

$$f) 5 + 7 + 9 + 11 + 13 + 15 = \sum_{j=2}^7 (2j + 1)$$

1.1.5.1.- Número de términos de una sumatoria: Considerando los límites inferior y superior, se puede determinar el número de términos que tendrá el desarrollo de una sumatoria simple finita.

Si “a” es el límite inferior y “b” el límite superior, el número “N” de términos estará dado por:

$$N = b - a + 1 \quad (1.2)$$

relación fácil de comprobar en los ejemplos anteriores.

1.1.5.2.- Propiedades operativas de las sumatorias: De acuerdo con la definición de sumatoria, se pueden deducir muchas propiedades que simplifican las operaciones que con ellas se realizan. A continuación, mencionaremos las más útiles, para lo cual consideraremos los siguientes valores de “X” e “Y”:

$$X_1 = 2; \quad X_2 = 3; \quad X_3 = 4$$

$$Y_1 = 1; \quad Y_2 = 5; \quad Y_3 = 6$$

1.1.5.2.1.- Primera propiedad:

$$\sum_{i=1}^n (X_i \pm Y_i) = \sum_{i=1}^n X_i \pm \sum_{i=1}^n Y_i$$

$$\text{Prueba: } \sum_{i=1}^3 (X_i + Y_i) = (X_1 + Y_1) + (X_2 + Y_2) + (X_3 + Y_3)$$

$$(2 + 1) + (3 + 5) + (4 + 6) = 3 + 8 + 10 = 21$$

por otro lado:

$$= X_1 + X_2 + X_3 = 2 + 3 + 4 = 9, \quad \text{y}$$

$$\sum_{i=1}^3 Y_i = Y_1 + Y_2 + Y_3 = 1 + 5 + 6 = 12$$

entonces, $21 = 9 + 12$, que es lo que queríamos demostrar (l.q.q.d.). Esto también es válido para la resta.

1.1.5.2.2.- Segunda propiedad: $\sum_{i=a}^b K = K(b - a + 1)$. En el caso especial en el que

$a = 1$, entonces: $\sum_{i=1}^b K = K \cdot b$

Prueba: $\sum_{i=2}^3 5 = 5 + 5 = 10$, al hacerlo sin aplicar la propiedad. Al aplicarla, resulta que:

$$= 5(3 - 2 + 1) = 5(2) = 10 \quad \text{l.q.q.d.}$$

1.1.5.2.3.- Tercera propiedad: $\sum_{i=1}^n K \cdot X_i = K \sum_{i=1}^n X_i$

Prueba: $\sum_{i=1}^3 5 \cdot X_i = 5X_1 + 5X_2 + 5X_3 = 5 \times 2 + 5 \times 3 + 5 \times 4 = 10 + 15 + 20 = 45$ por otro lado:

$$5 \sum_{i=1}^3 X_i = 5(X_1 + X_2 + X_3) = 5(2 + 3 + 4) = 5(9) = 45, \quad \text{l.q.q.d.}$$

1.1.5.2.4.- Cuarta propiedad: $\sum_{i=1}^n X_i^2 \neq (\sum_{i=1}^n X_i)^2$

$$\text{Prueba: } \sum_{i=1}^3 X_i^2 = X_1^2 + X_2^2 + X_3^2 = 2^2 + 3^2 + 4^2 = 4 + 9 + 16 = 29$$

por otro lado:

$$(X_1 + X_2 + X_3)^2 = (2 + 3 + 4)^2 = (9)^2 = 81$$

$29 \neq 81$ l.q.q.d.

1.1.5.2.5.- Quinta propiedad: $\sum_{i=1}^n X_i \cdot Y_i \neq \left(\sum_{i=1}^n X_i\right) \cdot \left(\sum_{i=1}^n Y_i\right)$

Prueba: $\sum_{i=1}^3 X_i \cdot Y_i = X_1 \cdot Y_1 + X_2 \cdot Y_2 + X_3 \cdot Y_3 = 2 \times 1 + 3 \times 5 + 4 \times 6 = 2 + 15 + 24 = 41$

por otro lado:

$$\left(\sum_{i=1}^3 X_i\right) \cdot \left(\sum_{i=1}^3 Y_i\right) = (X_1 + X_2 + X_3)(Y_1 + Y_2 + Y_3) = (2 + 3 + 4)(1 + 5 + 6) = (9)(12) = 108 \neq 41 \text{ l.q.q.d.}$$

1.1.5.2.6.- Sexta propiedad: $\sum_{i=1}^n \frac{1}{X_i} \neq \frac{1}{\sum_{i=1}^n X_i}$

Prueba: $\sum_{i=1}^3 \frac{1}{X_i} = \frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} = \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = 1,083$

por otro lado

$$\frac{1}{\sum_{i=1}^3 X_i} = \frac{1}{X_1 + X_2 + X_3} = \frac{1}{2 + 3 + 4} = \frac{1}{9} = 0,11$$

$1,083 \neq 0,11$ l.q.q.d.

1.2.- DEFINICIÓN DE ESTADÍSTICA

La *Estadística* es una herramienta que se utiliza en el método científico para hallar las regularidades del comportamiento colectivo. Estas regularidades sirven para describir lo que ocurre con los datos o para hacer predicciones. Por tanto, la Estadística estudia sólo fenómenos de masa y les sirve a los investigadores en cualquier campo de las ciencias: Física, Química, Biología, Economía, Sociología, Ingeniería, Medicina, Educación, etc.

La función principal de la *Estadística* es elaborar principios y métodos que nos ayuden a tomar decisiones frente a la incertidumbre. La mayoría de los libros e investigaciones especializadas contienen resultados basados en el análisis estadístico. Aquellos profesionales que no conozcan de Estadística, tendrán serias dificultades para ser expertos en sus respectivos campos científicos.

Las técnicas estadísticas se utilizan para lograr, generalmente, uno de los siguientes objetivos:

1.2.1.- Descripción de un gran número de datos: Reduciéndolos a un pequeño número de características que concentran la parte más importante y significativa de la información proporcionada por los datos. A este proceso se le conoce como “reducción de datos”. La descripción supone que los datos que vienen expresados tal y como fueron recogidos, deben ser organizados y presentados sistemáticamente en cuadros o tablas. De esta forma nace lo que se conoce como *Estadística Descriptiva*.

1.2.2.-Análisis de la población a través de muestras: Con el objeto de formarse un concepto acerca de una población o universo y adoptar decisiones, puede no ser necesario estudiar a todos y cada

uno de sus miembros sino solo a una parte de ellos. Esta parte puede ser bien una muestra o una subpoblación representativa de toda la población. La preocupación entonces, del análisis estadístico, es inferir propiedades para una población sobre la base de resultados muestrales conocidos. Para ello se han creado diferentes mecanismos estadísticos como la teoría del muestreo y el cálculo de las probabilidades. Estos aspectos corresponden esencialmente a la *Inferencia Estadística*.

1.2.3.- Predicción o comportamiento de los fenómenos en el futuro: Poder predecir o estimar el comportamiento de los fenómenos, constituye la máxima aspiración práctica de toda ciencia. Este objetivo de predicción y previsión está implícito tanto en la descripción como en el análisis estadístico, puesto que en general interesa orientar la toma de decisiones con vigencia y efecto en el futuro. “El pasado puede ser evaluado, el presente descrito y el futuro puede ser previsto”. La predicción puede entenderse como la estimación de resultados en el futuro.

Con este objetivo se pretende responder, por ejemplo, ¿Cuántos niños iniciarán el preescolar en la provincia de Valencia dentro de dos años? ¿Cuánto se estima que será la población en México dentro de cinco años?, etc. Naturalmente que la precisión en las estimaciones y proyecciones depende del grado de conocimiento que se tenga del comportamiento pasado y presente de las variables de estudio.

1.3.- TERMINOLOGÍA ESTADÍSTICA

Algunos términos que se utilizan dentro del campo estadístico son los siguientes:

1.3.1.- Población: Es el conjunto o colección de personas o cosas sobre los que se va a realizar un estudio estadístico.

1.3.2.- Elementos: Son cada una de las personas o cosas que forman parte de la población.

1.3.3.- Caracteres: Son las propiedades o cualidades que tienen cada uno de los elementos de la población. Es lo que se estudia estadísticamente en cada uno de los elementos. Por ejemplo, la procedencia, la edad, peso, estatura y color del cabello de un grupo de personas; el número de Kg. por hectárea obtenidos durante la cosecha en cada una de las fincas productoras de determinada región, etc.

1.3.3.1.- Variables: Los *caracteres* que pueden ser manifestados mediante números como la edad, estatura, ingreso económico mensual, suelen llamarse *variables* y vienen expresados mediante *valores*.

1.3.3.2.- Atributos: Los *caracteres cualitativos* o que se manifiestan mediante palabras como la procedencia, estado civil, sexo, suelen llamarse *atributos* y vienen expresados mediante *modalidades*.

A cada *valor* de una variable y a cada *modalidad* de un atributo se les llama un *dato*. De igual forma las variables pueden ser *continuas*, cuando admiten valores decimales (por ejemplo, el peso de una persona) y *discretas*, cuando no admiten valores decimales (por ejemplo, el número de hijos en una familia).

1.3.4.- Censo: Para efectuar el análisis estadístico de una población puede ser necesario estudiar todos y cada uno de los elementos que componen esa población. A esta recaudación exhaustiva de datos de *todos* los elementos de la población se le llama *censo*.

1.3.5.- Subpoblación: Puede ocurrir que para efectuar el estudio estadístico no sea necesario extraer datos de todos y cada uno de los elementos de la población. En ocasiones puede ser suficiente extraer datos sólo de un conjunto de elementos de la población. Cuando los elementos

seleccionados poseen ciertas características especiales que no se presentan en los restantes elementos de la población, se dice que tenemos una *subpoblación*. Como ejemplo, supóngase que, para conocer el índice académico promedio de los estudiantes de determinada carrera, pueda ser suficiente con conocer el índice académico de los estudiantes del sexo femenino. Observe que cada uno de los elementos que se han seleccionado para el estudio posee una característica que no se presenta en los restantes elementos y que es pertenecer al sexo femenino.

1.3.6.- Muestra: Cuando se desea extraer datos sólo de una parte de la población, pero donde los elementos seleccionados no posean ninguna característica que los diferencie de los elementos no seleccionados, se dice que se está tomando una muestra.

El tamaño de la *muestra* a tomar va a depender del grado de incertidumbre que se desea tener cuando se pretenda inferir los resultados obtenidos en la muestra hacia toda la población. Es obvio que, a mayor tamaño de muestra, menor grado de incertidumbre; de tal manera que, si en lugar de tomar una *muestra* se realiza un *censo*, el grado de incertidumbre en los resultados se reduce a cero.

Los métodos de *muestreo* se basan en la teoría de las probabilidades y existe una vasta información sobre el tema, en el cual, por cierto, se ha profundizado mucho en la estadística moderna.

Un ejemplo de *muestra* es el siguiente: “Para conocer el índice académico promedio de los estudiantes de determinada carrera, con un grado de confiabilidad del 95%, es decir con un grado de incertidumbre del 5%, puede ser suficiente con conocer el índice académico de sólo 5 estudiantes de los 50 que integran esa carrera. La selección de estos 5 estudiantes (elementos) que integrarán la muestra, se hará de tal forma que al final se notará que no presentan ninguna característica especial que los diferencie de los estudiantes que no fueron seleccionados.

1.3.6.1.- Muestreo aleatorio: Aquí los elementos de la muestra se seleccionan al azar. Entre las distintas formas en que puede realizarse este muestreo se encuentran el muestreo aleatorio simple y el muestro aleatorio estratificado.

1.3.6.1.1.- Muestreo aleatorio simple: Para llevar a cabo esta forma de muestreo, una vez que se ha calculado el número de elementos que integrarán la muestra de acuerdo al tamaño de la población y al grado de incertidumbre que se desea en los resultados, se realizará un sorteo entre todos los elementos que constituyen la población en estudio para conocer quiénes resultarán favorecidos y conformarán la muestra. Ejemplo: Supongamos que deseamos conocer la edad promedio de los 50 estudiantes de la sección 01 del curso de Estadística y que esta variable queremos inferirla utilizando una muestra que nos provea un valor con un 98% de confiabilidad. Al efectuar los cálculos de cuál debe ser el tamaño mínimo de esta muestra nos encontramos, que, por ejemplo, es de 10 alumnos.

Para conocer quiénes serán los 10 alumnos seleccionados para conformar esta muestra, que deseamos se tome mediante un muestreo aleatorio simple, se procede a colocar el nombre y apellido, cédula de identidad, DNI, o cualquiera otra identificación de cada uno de los 50 estudiantes en un listado de computadora o en un papel. Luego se le pide al computador que seleccione, o se escoge manualmente, al azar, 10 identificaciones de las 50. Estos 10 nombres seleccionados constituirán la muestra y a esos alumnos se les preguntará su edad; de tal forma, que la edad promedio de estos 10 alumnos, podemos decir con un 98% de probabilidad, que será la edad promedio de los 50 estudiantes de la sección 01.

1.3.6.1.2.- Muestreo aleatorio estratificado: Para obtener una *muestra aleatoria estratificada*, se divide previamente la población en diferentes estratos de acuerdo a las modalidades que pueda tener un atributo o a los valores que pueda tomar una variable. De esta forma resultan varios grupos

que son más homogéneos, de acuerdo a un carácter, que la población como un todo. Luego se realizan muestreos aleatorios simples en cada uno de los estratos y el tamaño de la “submuestra” en cada estrato, dependerá del tipo de muestreo estratificado del cual se trate. Veamos:

1.3.6.1.2.1.- Proporcional: En este caso la *proporción* que guarda cada una de las *sub-muestras* dentro de la muestra es igual a la proporción que guarda cada uno de los estratos en la población. Veámoslo con un ejemplo: Supóngase que para efectuar la investigación sobre la edad promedio de los 50 estudiantes mencionados en la sección 1.3.6.1.1, se divide a los estudiantes previamente en estratos de acuerdo al atributo “lugar de procedencia”, de la siguiente forma:

Lugar de procedencia	Nº de estudiantes
Región Llanera	20
Región andina	15
Región Central	10
Otras Regiones	5
Total	50

Supongamos que el tamaño de muestra ideal para efectuar este estudio, con un 98% de seguridad al inferir el resultado de la muestra a toda la población, es nuevamente 10 estudiantes. Entonces, para que el número de elementos de cada estrato en la muestra guarde la misma proporción que el número de elementos de cada estrato en la población, se procede de la siguiente forma:

Proporción existente entre el tamaño de la muestra y el de la población = $\frac{10}{50} = 0,20$

Aplicando esta proporción al número de estudiantes que posee cada estrato, se obtiene el número de estudiantes con el que contribuirá cada segmento para la conformación de la muestra. Veamos:

Lugar de procedencia	Nº de estudiantes en población	Nº de estudiantes en muestra
Región Llanera	20	4
Región andina	15	3
Región Central	10	2
Otras Regiones	5	1
Total	50	10

Posteriormente se procede a hacer un listado en el computador (o en papel) con las identidades de los 20 estudiantes procedentes de la Región Llanera y se seleccionan, al azar, 4 de estas identidades para integrar la muestra; se hace otro listado con las identidades de los 15 estudiantes procedentes de la Región Andina y se seleccionan, al azar, 3 de estas identidades para integrar la muestra; se hace lo propio con las identidades de los 10 estudiantes procedentes de la Región Central y se seleccionan, al azar, 2 de estas identidades y, finalmente, de Otras Regiones se selecciona 1. Luego se les pregunta a los 4 estudiantes procedentes de la Región Llanera, a los 3 estudiantes procedentes de la Región Andina, a los 2 estudiantes procedentes de la Región Central y al estudiante seleccionado al azar de Otras Regiones cuáles son sus edades. Se obtiene el promedio y podemos

inferir, con un 98% de probabilidad, que esta será la edad promedio de los estudiantes de toda la sección 01 de Estadística, analizada.

1.3.6.1.2.2.- No proporcional: En este caso el tamaño de cada uno de los estratos dentro de la muestra no guarda ninguna proporción con el tamaño de cada uno de los estratos dentro de la población. Principalmente puede ser de dos tipos:

- a) **No proporcional desigual:** En este caso lo importante es que la suma del número de elementos con el que contribuye cada estrato para conformar la muestra, sea igual al tamaño final de ésta, sin guardar ninguna relación con el tamaño que tenga cada estrato en la población.

Por ejemplo, en el ejercicio del punto 1.3.6.1.2.1, si la muestra se hubiera tomado mediante un muestreo aleatorio estratificado no proporcional desigual, uno de los resultados posibles es el siguiente

Lugar de procedencia	Nº de estudiantes en población	Nº de estudiantes en muestra
Región Llanera	20	3
Región andina	15	3
Región Central	10	2
Otras Regiones	5	2
Total	50	10

Luego se procede a hacer los listados en el computador por estrato para seleccionar las identidades de los 3 alumnos de la Región Llanera, los 3 de la Región Andina, los 2 de la Región Central y los 2 de Otras Regiones, de manera similar a como se expuso en el anterior punto 1.3.6.1.2.1.

- b) **No proporcional igual:** En este caso el número de elementos que conforman la muestra es igual para todos los estratos. Veamos el siguiente ejemplo:
Supóngase que con los alumnos del ejercicio del punto 1.3.6.1.2.1 se desea tomar una muestra aleatoria estratificada no proporcional igual, de 12 elementos. El siguiente sería el resultado (nótese que si la muestra fuera de 10 elementos no se podría tomar):

Lugar de procedencia	Nº de estudiantes en población	Nº de estudiantes en muestra
Región Llanera	20	3
Región andina	15	3
Región Central	10	3
Otras Regiones	5	3
Total	50	12

Posteriormente se procede de manera similar al punto 1.3.6.1.2.1.

1.3.7.- Registros o Estadísticas: Un *registro* es un archivo de datos donde compradores, vendedores, exportadores o cualquier institución pública o privada lleva una enumeración sistemática de sus operaciones. Los libros donde se registran los nacimientos y decesos en una prefectura, Registro

Civil o ayuntamiento, los libros donde se almacena la información sobre el rendimiento estudiantil de los alumnos a lo largo de su carrera, son estadísticas de nacimientos y muertes y de rendimiento estudiantil por alumno, respectivamente.

1.3.8.- Encuestas: Al procedimiento mediante el cual se recaudan datos a través de una muestra o una subpoblación se le llama *encuesta*. La encuesta se diferencia del censo en que es menor el número de elementos estudiados y son menores los caracteres considerados. La encuesta, además, requiere de testimonios orales o escritos de personas vivas.

1.3.9. Estadístico: Un *estadístico* es un indicador numérico que se calcula con los datos suministrados por los elementos de muestras. La media aritmética $\bar{X}$ y la desviación típica S son *estadísticos*, puesto que se calculan con información proveniente de muestras.

1.3.10.- Parámetro: Los *parámetros*, por el contrario, se calculan con información proveniente de poblaciones o inferida a poblaciones. La media aritmética μ y la desviación típica σ , son parámetros ya que han sido calculados para lo que ocurre en una población.

1.3.11.- Razonamiento deductivo: Va de *lo general a lo individual*. Generalmente se utiliza cuando se aplican leyes matemáticas, físicas, etc.

1.3.12.-Razonamiento Inductivo: Es cuando se va de *lo particular a lo general*. En estos casos generalmente se utilizan muestras.

1.4.- PASOS A SEGUIR EN UNA INVESTIGACIÓN ESTADÍSTICA

Lo primero que debe definirse cuando va a efectuarse una investigación estadística, son los *alcances en el espacio y el tiempo* de esa investigación. ¿Los *caracteres* a investigar cambian con el tiempo? Si lo hacen, ¿a qué período de tiempo nos remitiremos? ¿Estudiaremos lo que pasa con los caracteres durante un día, una semana, un mes, un año? Por otro lado, ¿haremos un *censo* o tomaremos una muestra o una subpoblación? ¿En qué lugar geográfico están los elementos? ¿Están en un aula de clases, en toda una escuela, en un barrio o urbanización, en una parroquia, un municipio, todo un estado, provincia o todo un país? ¿Para *cuándo* se quieren los resultados del estudio? ¿Para dentro de unas horas, unos días, unas semanas o unos meses?

Todas estas preguntas deben realizárselas el investigador estadístico antes de comenzar la investigación, porque dependiendo de las respuestas, él utilizará una metodología u otra, tanto para recabar los datos, como para analizarlos y la investigación requerirá de unos recursos u otros, lo cual indudablemente cambiará el costo final.

Una vez que se han definido con claridad los tópicos anteriores, el investigador procederá a cumplir con los siguientes pasos:

1.4.1.- La recolección de los datos: La *recolección o recopilación* de datos es el momento en el cual el investigador se pone en contacto con los objetos o elementos sometidos a estudio con el propósito de obtener los datos o respuestas de los atributos o variables consideradas. A partir de estos datos se prepara la información estadística y se calculan medidas de resumen e indicadores para el análisis estadístico.

El trabajo de recolección de datos, en general, se puede realizar de dos maneras:

- a) La técnica de *investigación documental* o bibliográfica.
- b) La técnica de trabajo de campo.

Por su parte el trabajo de campo puede realizarse de dos formas:

b.1) *La observación y la exploración en el terreno*, que consiste en el contacto directo del investigador con el objeto de estudio.

b.2) *La encuesta*, que consiste, como ya dijimos anteriormente, en testimonios orales o escritos de personas vivas.

La fase de recolección de datos es uno de los puntos neurálgicos de la investigación. En consecuencia, deben existir todas las garantías de que los datos tomados son confiables y comparables y que no hay desviaciones, ni falta de representatividad.

a) La técnica de investigación documental: Mediante esta técnica se recopila o busca los datos en documentos o fuentes escritas y fílmicas de todo tipo. Estos datos pueden ser *internos*, lo cual significa que pertenecen a los archivos de la propia organización que hace el estudio, o *externos*, que se encuentran en otras fuentes. La información bibliográfica tanto interna, como externa, puede recabarse en libros, actas (que son documentos que recogen lo ocurrido en el momento mismo en que está pasando), videos e informes (documento escrito después de ocurrido el acontecimiento).

b) La técnica de trabajo de campo

b.1) La observación y exploración en el terreno: En este caso el investigador observa directamente en el terreno los casos o individuos en los que se produce un fenómeno, entrando en contacto con ellos. Requiere de personal competente y especializado el cual deberá actuar en forma objetiva y sin modificar los datos. Las pruebas químicas, biológicas y físicas que se efectúan en los laboratorios son ejemplos de este tipo de recolección de datos estadísticos. Las investigaciones policiales que realizan los cuerpos de seguridad del Estado, son otro ejemplo.

b.2) La encuesta: Como ya se ha dicho este medio de recolectar datos se aplica exclusivamente a personas. Existen diferentes medios para aplicar una encuesta, cada uno con sus ventajas y desventajas. Dependiendo de la rapidez con la cual se quieran obtener los datos y de la cantidad de recursos económicos con los que se cuente, puede utilizarse uno u otro de los siguientes medios:

b.2.1) La entrevista personal: La *entrevista personal* tiene la ventaja que el entrevistado y el entrevistador se encuentran frente a frente y por tanto existen garantías de que se responderán todas o casi todas las preguntas del cuestionario, el entrevistado entenderá qué significa cada pregunta y, adicionalmente, regresarán contestados los cuestionarios con relativa rapidez. Pero se trata de un procedimiento sumamente costoso: deberá reclutarse y entrenarse personal idóneo al cual deberá cancelarse todos los emolumentos que estipula la ley.

b.2.2) Por medio del teléfono: Este mecanismo para aplicar el cuestionario es sumamente rápido y relativamente económico. Basta colocar a uno o varios telefonistas los cuales llamarán a los elementos seleccionados y les aplicarán la encuesta. Tiene la desventaja que las personas que deben ser entrevistadas en algunas investigaciones, no todas puedan contar con teléfono.

b.2.3) Por medio del correo: El *correo* es un medio bastante barato y teóricamente llega a todas partes. Sin embargo, no hay garantía que los elementos seleccionados entiendan las preguntas del cuestionario, las respondan y lo regresen.

b.2.4) Por correo electrónico: El realizar encuestas a través de e-mail es bastante económico, pero tiene la desventaja, al igual que con el teléfono, que algunos o todos los elementos a encuestar

podieran no tener correo electrónico, Además, aún en el caso que lo tuvieran, no hay garantía de reenvío del cuestionario contestado.

1.4.1.1.- Preparación del cuestionario: El *cuestionario* es un instrumento constituido por un conjunto de preguntas sistemáticamente elaboradas, que se formulan al encuestado o

entrevistado, con el propósito de obtener los datos referentes a los caracteres considerados en el estudio. Permite la *descripción* de las características de individuos o grupos en relación con variables y atributos como el sexo, edad, años de educación, ocupación, ingresos, estado civil, vivienda, afiliación política, etc., y la *medición* de rasgos individuales o grupales, particularmente actitudes como la distancia social, percepción de la cohesión grupal, grado de prejuicio racial, permisividad sexual, religiosidad, ansiedad, alienación, etc.

La elaboración del cuestionario requiere de un conocimiento previo del fenómeno y tener precisado qué aspectos interesa estudiar. El cuestionario debe ser adaptado a las necesidades de la investigación y a las características del grupo que se estudia. De allí que su estructura y su forma deben ser cuidadosamente elaborados. La redacción de las preguntas requiere de experiencia y para comprobar su comprensión, es recomendable someter a prueba todo tipo de cuestionario.

Criterios para preparar el cuestionario: Para preparar adecuadamente un cuestionario debe tenerse presente lo siguiente:

- a) *Objetivos* de la investigación
- b) Tipo de caracteres a investigar
- c) *Características* del encuestado
- d) *Tiempo disponible* para recoger los datos
- e) *Técnica* de recolección

Características formales del cuestionario: A fin de facilitar su comprensión, en todo cuestionario debe considerarse:

- a) Su forma y tamaño
- b) La *calidad* del papel del formulario
- c) El tipo y color de la impresión

Formas y clases de preguntas que debe contener:

- a) *Preguntas abiertas:* Son aquellas en las que el interrogado responde con su propio vocabulario sin que se le ponga ninguna alternativa. Ejemplos: 1.- “Describa brevemente las tareas que realiza en el cargo que ocupa”. 2.- “Anote las revistas especializadas en temas económicos que Ud. conoce”.
- b) *Preguntas cerradas dicotómicas:* Son aquellas que sólo pueden responderse con un “sí” o un “no”, o simplemente cuando sólo tienen dos alternativas. Ejemplos: - ¿Sabe leer?
Si () 1 No () 2

Sexo M () 1 F () 2

¿Trabaja? Si () 1 No () 2

- c) *Preguntas cerradas de selección múltiple*: En este caso se proponen un conjunto de alternativas para la respuesta. Ejemplos:

Categoría docente: Instructor	() 1	- Edad: Menor de 20 años	() 1
Asistente	() 2	De 20 a 29 años	() 2
Agregado	() 3	De 30 a 39 años	() 3
Asociado	() 4	De 40 a 49 años	() 4
Titular	() 5	De 50 años y más	() 5

- d) *Preguntas literales*: Son preguntas abiertas cuyas respuestas se expresan con una palabra o cantidad. Ejemplos:

¿Cuál es su estado civil? ----- - ¿Cuál es su ocupación actual? -----

¿Cuál es su ingreso económico mensual? -----

- d) *Preguntas con respuestas en grado de intensidad*: Aquí las respuestas indican un grado de intensidad dentro de una escala creciente o decreciente y el encuestado tiene que elegir una de las posibilidades. Ejemplo:

La gestión del Rector actual en la Universidad, es:

Buena () Regular () mala () No opina ()

Momento de la codificación:

Pre-codificados: por ejemplo: SEXO 1) Masculino	ESTADO CIVIL 1) S
2) Femenino	2) C
	3) V
	4) D

Post-codificados

Pautas a considerar para la redacción y el contenido de las preguntas:

- Incluir sólo *preguntas estrictamente necesarias* para: 1.- Satisfacer objetivos; 2.- establecer controles de consistencia en las respuestas; 3.- Individualizar al informante.
- Incluir *preguntas* para que el informante recuerde.
- Evitar *preguntas* que obliguen al informante a realizar cálculos y pueda cometer errores.
- Evitar *palabras* que provoquen diferentes interpretaciones.
- Redactar en forma *clara y directa*.
- Las preguntas no deben sugerir respuestas.

- g) Evitar en lo posible preguntas íntimas.
- h) Colocar primero las *preguntas más simples* y menos íntimas.
- i) Cada pregunta debe referirse a *un solo aspecto* o asunto
- j) *No colocar un orden siempre igual* en las alternativas de respuesta de las preguntas
- k) Las preguntas *no deben apoyarse* en evidencias comprobadas
- l) Los *cambios repentinos en los temas deben evitarse*, colocando siempre preguntas de transición

1.4.2.- La organización de los datos: Como resultado de la recopilación se dispone de una cantidad de datos, registrados en los formularios tal y como fueron captados, en su forma más natural y estadísticamente desordenados. Para facilitar el análisis estadístico y la interpretación de los resultados es necesario organizar los datos.

Los siguientes son los pasos necesarios para organizar los datos:

1.4.2.1.- Corrección: Generalmente en los resultados de una encuesta se corrigen:

- a) Las respuestas incongruentes (chocan entre si varias respuestas).
- b) Escritura ilegible.
- c) Respuestas incompletas.
- d) Cálculos errados u omitidos.

1.4.2.2.- Clasificación: Los datos una vez corregidos son clasificados de acuerdo a uno o varios criterios, por ejemplo: de acuerdo a un criterio cronológico (como días, semanas, meses, etc.), o de acuerdo a un criterio geográfico (ubicación o lugar), o de acuerdo a criterios cuantitativos o cualitativos (edades, estaturas, estado civil).

1.4.2.3.- Registro: Luego que los datos son clasificados deben ser guardados en esa forma mediante algún procedimiento, bien sea manual o preferiblemente electrónico.

Una vez registrados los datos, los formularios o cuestionarios respondidos pueden ser destruidos.

1.4.3.- La presentación de los datos: Luego que los datos han sido registrados, se hace necesaria su presentación o conocimiento público de manera ordenada y entendible para cualquier persona, conozca o no, del tema que se investiga. Junto con los datos propiamente dichos, debe suministrarse información sobre de qué datos se trata, cómo se les ha clasificado, cuándo ocurrieron esos datos y en qué lugar se les tomó.

Existen tres maneras de *presentar los datos* dependiendo del volumen de éstos y del dinamismo que presenten. Estas tres maneras son:

1.4.3.1.- Mediante enunciados: Se produce la presentación de datos *mediante enunciados* o palabras cuando los datos recopilados presentan poca variación entre ellos. Un ejemplo de esta presentación se produce cuando un organismo policial presenta un informe sobre la investigación de un atraco.

1.4.3.2.- Mediante tablas estadísticas: Las *tablas o cuadros estadísticos* se construyen para visualizar, a modo de conjunto, a un gran volumen de datos. En una investigación estadística no siempre puede elaborarse una sola tabla que muestre toda la información recabada; generalmente deben construirse dos o más tablas, dependiendo de la forma como fueron clasificados los datos.

En vista que la construcción de tablas debe ser de esencial conocimiento para cualquier estudiante de estadística y que existen diversos criterios para construirlas dependiendo del número de elementos en estudio y del número y tipo de caracteres analizados, dedicaremos el II capítulo de este libro exclusivamente a este tema.

1.4.3.3.- Mediante gráficas estadísticas: Las *gráficas o figuras estadísticas* se construyen cuando se hace necesario observar el comportamiento de los datos al variar un carácter de estudio. Así, las gráficas poseen un ingrediente dinámico en su concepción intrínseca, que las convierte en una herramienta que comunica al investigador una visión de conjunto de los datos en ellas plasmados.

Dada la relevante importancia de las gráficas en la Estadística, se dedicará a ellas todo el capítulo III de este libro.

1.4.4.- El análisis de los datos: Los datos que generalmente se analizan estadísticamente son los datos de las tablas. El *análisis* se realiza para obtener indicadores, estadísticos o parámetros que nos permitan sacar conclusiones sobre el comportamiento de los datos recogidos. El *análisis estadístico* se realiza, entre otras, a través de la Estadística Descriptiva y la Estadística Inductiva.

Por su importancia dedicaremos varios capítulos de este libro al estudio del análisis estadístico y los principios que lo rigen.

1.4.5.- La Interpretación de los resultados: Una vez que se obtienen los indicadores a través del análisis estadístico, éstos deben ser interpretados por un experto en la materia objeto de la investigación, de tal forma que se pueda realizar la toma de decisiones que se desea, con un mínimo de riesgos.

1.5.- EL MÉTODO CIENTÍFICO

Para la realización de una investigación estadística se utiliza el *método científico*. Esto se manifiesta porque se cumplen, de una u otra forma, las siguientes premisas:

- Se basa en hechos observados
- Se formulan *hipótesis*
- Se realizan uno o varios *experimentos*
- Se obtienen *resultados* y se debe realizar su *interpretación*
- Se producen conclusiones y recomendaciones

1.5.1.- Características de un Experimento bien planeado

- Simplicidad
- Grado de *precisión*
- Ausencia de error sistemático
- Rango de validez de las *conclusiones*
- Cálculo del grado de *incertidumbre*

1.5.2.- Procedimiento para la Experimentación:

- Definición del *Problema*
- Determinación de los *objetivos*
- *Análisis crítico* del problema y de los objetivos
- Selección de los *tratamientos* a aplicar
- Selección del material experimental
- Selección del Diseño Experimental
- Selección de la unidad de observación y del número de repeticiones
- Consideración acerca de los *datos* que se van *a recabar*
- Esbozo del *análisis estadístico* y del resumen de los *resultados*
- Conducción del experimento
- Análisis de los datos e interpretación de los resultados
- Elaboración de un completo, legible y correcto *informe de la investigación*

EJERCICIOS PROPUESTOS

1.- Con la información presente en el siguiente cuadro:

Cuadro N° 1

Estudiantes de la UNELLEZ – San Carlos, clasificados de acuerdo a su sexo y la carrera que cursan. San Carlos, enero 2018.

SEXO	INGENIERÍA	TECNOLOG.	DERECHO	EDUCACIÓN
Masculino	380	150	330	100
Femenino	320	200	600	250

Fuente: Datos supuestos

- Calcule la razón de estudiantes de Derecho a estudiantes de Ingeniería.
- Calcule la razón de estudiantes del sexo masculino de tecnología a estudiantes del sexo femenino de Educación.
- Calcule la razón de estudiantes del sexo femenino a estudiantes del sexo masculino.

2.- Con los datos del Cuadro N°1, calcule:

- La proporción de los estudiantes de Derecho de la UNELLEZ – San Carlos.
- La proporción por sexo dentro de los estudiantes de Tecnología.

3.- Con los datos del Cuadro N°1, calcule:

- El porcentaje de los estudiantes de la UNELLEZ – San Carlos que estudian Ingeniería
- El porcentaje de los estudiantes de la UNELLEZ – San Carlos que estudian Educación.

4.- Si en toda la ciudad de San Carlos en el año 2018 vivían 320.000 personas, y el número de estudiantes de la UNELLEZ – San Carlos era el que muestra el Cuadro N° 1, diga:

- ¿Qué tasa de estudiantes unelliztas tenía San Carlos en el año 2018?
- ¿Qué tasa de estudiantes de Ingeniería en la UNELLEZ tenía San Carlos en el año 2018?

5.- Desarrolle en forma completa cada una de las siguientes sumatorias:

$$a) \sum_{i=6}^{10} W_i^2 \quad ; b) \sum_{h=2}^4 (X_h + h) \quad ; c) \sum_{j=1}^5 3(V_j - 2) \quad ; d) \sum_{j=0}^3 \left(\frac{X_j^i Y_z^i}{K X_i Y_j^j} - 3^2 \right) \quad ; e) \sum_{i=3}^7 5(X_{2i} + 1)$$

6.- Calcule las siguientes sumatorias: Si $X_1 = 4$; $X_2 = -3$; $X_3 = 6$ y $X_4 = -1$;

$$a) \sum_{i=1}^4 X_i^2 (X_i - 3); \quad b) \sum_{i=1}^3 (X_i + 1); \quad c) \sum_{i=2}^4 \frac{(X_i + 2)}{X_i}$$

10.- Los hoteles de una determinada población han sido agrupados de acuerdo a su comodidad y los servicios que prestan a sus huéspedes de acuerdo a la siguiente manera:

CATEGORÍA	Nº DE HOTELES
Económica	30
Estándar	10
Lujo	3

Con esta información:

- a) Tome una subpoblación de 5 elementos.
- b) Tome una muestra aleatoria simple de 8 elementos.
- c) Tome una muestra aleatoria estratificada proporcional de 6 elementos.

11.- Explique, de la manera más sucinta posible, cómo aplicaría todos los pasos necesarios para realizar las siguientes investigaciones estadísticas:

- a) Se desea conocer cuál es la edad promedio de todos los estudiantes de la UNELLEZ.
- b) Se desea conocer cuál es el candidato presidencial que actualmente tiene mayor aceptación en la provincia de Portuguesa.

