



Análisis Estadístico Univariado

ING. LORENZO JOVANNY CEVALLOS TORRES MSC.
ING. NELLY AMERICA VALENCIA MARTINEZ MSC.
DR. ROOSEVELT LINDORFO BARROS MORALES MSC



VICERRECTORADO DE
INVESTIGACIÓN
GESTIÓN DEL
CONOCIMIENTO
Y POSGRADO

compAs

Análisis Estadístico Univariado

Autores:

ING. LORENZO JOVANNY CEVALLOS TORRES MSC.
ING. NELLY AMERICA VALENCIA MARTINEZ MSC.
DR. ROOSEVELT LINDORFO BARROS MORALES MSC





VICERRECTORADO DE
INVESTIGACIÓN
GESTIÓN DEL
CONOCIMIENTO
Y POSGRADO



Análisis Estadístico Univariado

Autores: ING. LORENZO JOVANNY CEVALLOS TORRES MSC.
ING. NELLY AMERICA VALENCIA MARTINEZ MSC.
DR. ROOSEVELT LINDORFO BARROS MORALES MSC

Primera edición: marzo 2017

Diseño de portada y diagramación:

Grupo Compás
Equipo Editorial

Dirección de Investigación
Vicerrectorado de Investigación, Gestión del Conocimiento y Posgrado
Universidad de Guayaquil

ISBN: 78-9978-59-143-7

Quedan rigurosamente prohibidas, bajo las sanciones en las leyes, la producción o almacenamiento total o parcial de la presente publicación, incluyendo el diseño de la portada, así como la transmisión de la misma por cualquiera de sus medios, tanto si es electrónico, como químico, mecánico, óptico, de grabación o bien de fotocopia, sin la autorización de los titulares del copyright.

AGRADECIMIETOS

A Mi Padre Celestial, por haberme dado la inteligencia y haberme guiado espiritualmente para poder redactar este libro.

A mi esposa Rossy y a mis tres hijas, Nicolle, Doménica y Elianai, por su paciencia y amor, por su sacrificio y comprensión, LAS AMO.

A mis estudiantes que poco a poco fueron dando de su aporte para poder Completar este libro.

A la Universidad de Guayaquil, y a la Facultad de Filosofía y letras de la Educación, por su empeño en el lanzamiento de este libro
(LORENZO CEVALLOS)

A mi compañero Lorenzo Cevallos por permitirme contribuir y ser parte del proceso de este libro.

A mi esposo Hans por su ayuda incondicional en la realización de todos los trabajos emprendidos.

A mi familia por ser mi fuente de apoyo constante en toda mi vida.
(Nelly)

El Agradecimiento sincero a Nelly y Lorenzo docentes respetables de la U.G. A mi esposa ANITA que desde la eternidad me ilumina y a Morayma - Carlos hijos incondicionales que apoyan en todos los proyectos propuestos en los caminos de la vida. **(Roosevelt)**

DEDICATORIA

A mi estudiante el Ing. **ALFREDO YIXON CASTRO BERNITA**, quien en vida me ayudo a redactar este libro y concretar este desafío, de poder compartir mis experiencias y sobre todo mi conocimiento a la comunidad universitaria de la Universidad de Guayaquil... descansa en paz. **(Lorenzo)**

A mis hijas **Cynthia** y **Emely**, a ellas que son la razón de vivir y luchar; que este esfuerzo mío les sirva de ejemplo. **(Nelly)**

Una dedicatoria especial a los compañeros docentes de la U. G, Facultad de Filosofía Letras y Ciencias de la Educación como también a todos los Estudiantes con quienes compartimos procesos de interaprendizaje. **(Roosevelt)**

INDICE GENERAL DE CONTENIDOS

PRÓLOGO	11
1. EL UNIVERSO UNIVARIADO Y SUS APLICACIONES	14
1.1 Introducción de la Estadística	14
1.2 ¿Qué es la Estadística?	14
1.3 ¿Para qué me sirve la estadística?	15
1.4 División de la Estadística	15
1.4.1 Estadística Descriptiva	15
1.4.2 Estadística Inferencial	16
1.5 DIFERENCIA ENTRE POBLACIÓN, UNIDAD DE INVESTIGACIÓN, MUESTRA, OBSERVACIÓN	17
1.5.1 Población objetivo.....	17
1.5.2 Unidad de investigación.....	17
1.5.3 Muestra	18
1.5.4 Observación	18
1.6 PARÁMETRO Y ESTIMADOR	19
1.6.1 Parámetro	19
1.6.2 Estimador o estadístico.....	19
1.7 TIPOS DE DATOS Y ESCALAS DE MEDICIÓN - VARIABLES	19
1.7.1 Variables Cuantitativas.....	20

1.7.2	Variables Cuantitativas Discretas	20
1.7.3	Variables Cuantitativas Continuas	21
1.8	Variables Cualitativas	22
1.8.1	Variables Cualitativas Nominales	22
1.8.2	Variables Cualitativas Ordinales.....	22
1.8.3	ANÁLISIS ESTADÍSTICO	22
1.8.3.1	Tablas de Frecuencia	23
1.8.3.2	Clase	24
1.8.3.3	Marca de clase	24
1.8.3.4	Frecuencia Absoluta	24
1.8.3.5	Frecuencia Relativa	24
1.8.3.6	Frecuencia Acumulada	25
1.8.3.7	Frecuencia Relativa Acumulada	25
1.8.4	ANÁLISIS DE DATOS CUANTITATIVOS	27
1.8.4.1	Tabla de Frecuencia por Intervalo	27
1.8.4.2	CONSTRUCCIÓN DE UNA TABLA DE FRECUENCIAS CUANTITATIVA	28
1.8.4.3	¿Cómo saber cuántos intervalos considerar?	28
1.8.4.4	Calculo del número de intervalos. –	28
1.8.4.5	Calculo del ancho del intervalo. –	29
2.	GRÁFICOS ESTADÍSTICOS PARA UNA DISTRIBUCIÓN DE FRECUENCIAS CUANTITATIVA Y CUALITATIVA.....	34
2.1	Diagrama de Barras	35

2.2	Histograma de frecuencia	35
2.2.1	Polígono de frecuencias	36
2.2.2	Ojiva	37
2.2.3	Diagrama de Pastel	37
2.2.4	Diagrama de Cajas.....	42
3.	MEDIDAS DE TENDENCIA CENTRAL Y DISPERSIÓN (DATOS NO AGRUPADOS)	47
3.1	Definición de estudio:	48
3.2	Media Muestral.....	49
3.3	Mediana	49
3.4	Moda	50
4.	MEDIDAS DE POSICIONAMIENTO	52
4.1	CUANTILES.....	53
4.2	PERCENTILES, ORDEN I.....	54
4.3	DECILES.....	55
4.4	CUARTILES	56
5.	MEDIDAS DE DISPERSIÓN O VARIABILIDAD	64
5.1	El Rango.....	64
5.1.1	Varianza Muestral	64
5.2	Desviación Estándar	65
5.3	Coficiente de Variación (C.V)	65
6	MEDIDAS DE FORMA.....	72
6.1	Coficiente de Asimetría (g1).....	72

6.2	Coeficiente de Curtósís o Aplastamiento (g2)	76
7	DISEÑO MUESTRAL	79
7.1	Definición de Muestreo	79
7.2	Tamaño de la Muestra	79
7.2.1	Población Finita	79
7.2.2	Población Infinita	82
7.3	Tipos de Muestreos: Muestreos no Probabilísticos....	83
7.3.1	Muestreo por Cuotas	84
7.3.2	Muestreo de Opinión o Intencional	85
7.3.3	Muestreo Casual o Incidental	86
7.3.4	Muestreo Bola de Nieve.....	87
7.3.5	Muestreo Discrecional	88
7.3.6	Muestreo Aleatorio Simple	90
7.3.7	Muestreo Estratificado.....	92
7.3.8	Muestreo por Conglomerados	94
7.3.9	Muestreo Sistemático	95
7.4	Elaboración del Cuestionario	97
7.4.1	Preguntas Abiertas	98
7.4.2	Preguntas Cerradas	99
7.4.3	Preguntas Dicotómicas y excluyentes.....	100
7.4.4	Preguntas de Respuesta Múltiple y Excluyente....	100
7.4.5	Preguntas No excluyentes.....	100
7.5	Elaboración del Informe	100

7.5.1	Título:	101
7.5.2	Resumen:.....	101
7.5.3	Introducción:.....	101
7.5.4	Métodos:	101
7.5.5	Resultados:.....	101
7.5.6	Análisis:	101
7.5.7	Conclusiones	101
8	“ANÁLISIS ESTADÍSTICO PARA DETERMINAR LA PREVALENCIA DEL PLAGIO ACADÉMICO EN TRABAJOS DE INVESTIGACIÓN CIENTÍFICA”	106
8.1	DISEÑO DE LA INVESTIGACIÓN	107
8.2	POBLACION Y MUESTRA	107
8.2.1	POBLACIÓN OBJETIVO.....	107
8.2.2	ELEMENTOS QUE FORMAN PARTE DEL CÁLCULO DEL TAMAÑO DE LA MUESTRA	109
8.2.3	CÁLCULO DEL TAMAÑO DE LA MUESTRA	109
8.3	INSTRUMENTOS DE RECOLECCION DE DATOS.....	110
8.3.1	ENCUESTA	111
8.3.2	TABULACIONES DE LA ENCUESTA	114
8.3.3	DESCRIPCION Y TIPO DE CADA UNA DE LAS VARIABLES	120
8.4	ANÁLISIS ESTADÍSTICO UNIVARIADO	121
8.4.1	Variable, Género del entrevistado	122
8.4.2	Variable. Edad del entrevistado (en años).	124

8.4.3	Variable: “Indique el nombre de la Institución de Educación Superior (IES) a la que Pertenece”.	131
8.4.4	Variable: Carrera a la que pertenece el entrevistado	133
8.4.5	Variable: Horas semanales dedica para hacer trabajos de investigación	135
8.4.6	Variable: Cuando finaliza su trabajo de investigación ¿Cree usted que adquirió los conocimientos adecuados acerca del tema que investiga?	137
9	ANÁLISIS ESTADÍSTICO DESCRIPTIVO APLICADO A LA DESERCIÓN ESTUDIANTIL DE LOS CURSOS DE NIVELACIÓN, CASO DE ESTUDIO: UNIVERSIDAD DE GUAYAQUIL.....	140
9.1	Población Objetivo.	140
9.2	Marco Muestral.....	141
9.3	Determinación del tamaño de la muestra.	142
9.4	Instrumento de Recolección de Datos.....	142
9.5	Diseño del cuestionario.....	143
9.5.1	ENCUESTA	144
9.6	DESCRIPCIÓN Y CODIFICACIÓN DE VARIABLES	147
9.6.1	CARACTERÍSTICAS GENERALES DEL ENTREVISTADO	147
9.6.2	SITUACIÓN LABORAL DEL ESTUDIANTE	151
9.6.3	INFORMACION FAMILIAR.....	152
9.6.4	INFORMACION ACADEMICA	155

9.6.5	INFORMACION RESPECTO A LA CAUSA DE DESERCIÓN	158
9.7	Metodologías estadísticas para identificar los factores asociados a la deserción estudiantil.	161
9.8	INTERPRETACIÓN Y ANÁLISIS DE LOS RESULTADOS	161
9.8.1	Análisis Univariado de los Datos	161
9.8.2	SECCIÓN: INFORMACIÓN GENERAL DEL ENTREVISTADO.....	162
9.8.2.1	Variable: Facultad donde estudia el entrevistado	162
9.8.2.2	Variable: Género del Entrevistado	163
9.8.2.3	Variable: Edad del entrevistado	164
9.8.2.4	Variable: Estado Civil del entrevistado	166
9.8.2.5	Variable: Tipo Colegio	167
9.8.2.6	Variable: Jornada de Estudio.....	168
9.8.2.7	Variable: Tipo Bachillerato	169
9.8.3	Sección: INFORMACIÓN FAMILIAR	170
9.8.3.1	Variable: Tiempo Dedicación Trabajo	170
9.8.3.2	Variable: Nivel Educativo Del Padre.....	171
9.8.3.3	Variable: Nivel Educativo De la Madre	172
10	ANÁLISIS DE DATOS CUALITATIVOS	174
10.1	CONSTRUCCIÓN DE UNA TABLA DE CONTINGENCIA	174

10.2	CASO DE ESTUDIO: FACTORES QUE INCIDEN EN EL RENDIMIENTO ACADÉMICO EN LOS ESTUDIANTES DE NIVEL SUPERIOR CON RESPECTO A SU ESTATUS LABORAL.....	175
10.2.1	Definición de las Variables Cualitativas que serán objeto de este estudio.	176
10.3	ANALISIS DE TABLAS DE CONTINGENCIA	178
	CONCLUSIONES	189
	Bibliografía	191

PRÓLOGO

Muchas veces el solo mencionar la palabra estadística llega inmediatamente a nuestra mente imágenes de grandes números agrupados, así como grandes arreglos y tablas, y situaciones relacionadas con nacimientos, muertes, impuestos, poblaciones, ingresos, deudas, crédito, etc. hasta llenarnos de un profundo sentimientos de querer no saber de ella por lo complicado que resulta.

Pero en realidad la Estadística es mucho más que sólo números, gráficas bonitas y formulas engorrosas. En realidad, la estadística es una ciencia tan antigua que data aproximadamente hacia el año 3050 antes de Cristo y fueron utilizados por los egipcios, pero en realidad quienes más le dieron provecho, fueron los romanos, quienes pudieron emplear mejor los recursos de la estadística, es así que cada cinco años realizaban un censo poblacional, de tal manera que tenían el control de los nacimientos, defunciones y matrimonios. Además, debemos considerar que la estadística es por sí misma auxiliar de todas las demás ciencias, tales como, la medicina, la ingeniería, la sociología, la psicología, también es muy utilizada en el ámbito de los gobiernos, los mercados, entre otras. Es importante considerar que la ausencia de la estadística nos conllevaría a una situación bastante desfavorable, dejando a los administradores y ejecutivos sin información vital a la hora de tomar decisiones en tiempos de incertidumbre.

Para una mejor aplicación y uso, este libro está diseñado a través de casos reales de estudio, es decir, está conformado por un conjunto de materiales didácticos y tecnológicos, en pos de cumplir con un objetivo medular que busca la integración de los contenidos de la estadística descriptiva. Como objetivos derivados, este proyecto pretende incorporar el uso de medios tecnológicos, para la mejor resolución de problemas y valorar el

uso de un lenguaje preciso y técnico, como organizador de las habilidades y destrezas individuales y/o grupales.

La aplicación de algunos proyectos durante algunos ciclos académicos, ha permitido tener algunos resultados preliminares, el análisis de alcances, limitaciones y la recopilación de la opinión de los estudiantes, lo que posibilita esbozar algunas conclusiones. La propuesta promueve a la integración de contenidos estadísticos y la utilización de las tecnologías informáticas y de comunicación, destinados a fomentar el intercambio de conocimientos entre alumnos, el trabajo colaborativo y significativo.

PARTE 1

**Fundamentos
teóricos**

1. EL UNIVERSO UNIVARIADO Y SUS APLICACIONES

2. Introducción de la Estadística

La Estadística es la Ciencia de la sistematización, recogida, ordenación y presentación de los datos referentes a un fenómeno que presenta variabilidad o incertidumbre para su estudio metódico, con objeto de deducir las leyes que rigen esos fenómenos y poder de esa forma hacer previsiones sobre los mismos, tomar decisiones u obtener conclusiones



3. ¿Qué es la Estadística?

La Estadística es la ciencia donde se obtienen, analizan, procesan y muestran datos referentes a un fenómeno que presenta

variabilidad o incertidumbre para su estudio metódico, con objeto de deducir las características de una población objetivo, y poder de esa forma actuar sobre los mismos, tomando decisiones u obtener conclusiones. (*Depool, Monasterio, 2013*) nos indica: “La Estadística es una ciencia cuyo método consiste en recopilar, presentar, analizar e interpretar datos numéricos extraídos de hechos reales e inferir de ellos, conclusiones lógicamente aceptables”.

4. ¿Para qué me sirve la estadística?

En realidad, la estadística se ocupa en general de los fenómenos observables, se utiliza como tecnología al servicio de la ciencia donde la variabilidad y la incertidumbre forman parte de la naturaleza. (*Freund, 2004*) indica, “Es el estudio de los mejores modos de acumular y analizar datos y establecer conclusiones acerca del coleccionativo del que se han recogido tales datos”

5. División de la Estadística

La Estadística para su mejor estudio se ha dividido en dos grandes ramas: la Estadística Descriptiva y la Inferencial.

6. Estadística Descriptiva

Consiste sobre todo en la presentación de datos en forma de tablas y gráficas. Esta comprende cualquier actividad relacionada con los datos y está diseñada para resumir o describir los mismos sin factores pertinentes adicionales; esto es, sin intentar inferir nada que vaya más allá de los datos, simplemente los describe de manera cuantitativa, como lo indica *Delgado R. (2004)*. “La estadística

descriptiva se encarga de recolectar, ordenar, analizar y representar los datos”.



Ilustración 1: Estadística Descriptiva

7. Estadística Inferencial

Se deriva de muestras, de observaciones hechas sólo acerca de una parte de un conjunto numeroso de elementos y esto implica que su análisis requiere de generalizaciones que van más allá de los datos. La Estadística Inferencial investiga o analiza una población partiendo de una muestra tomada. Como lo indica Gómez M. (2005).” La inferencia estadística se puede escoger varias alternativas o formas de acción”.

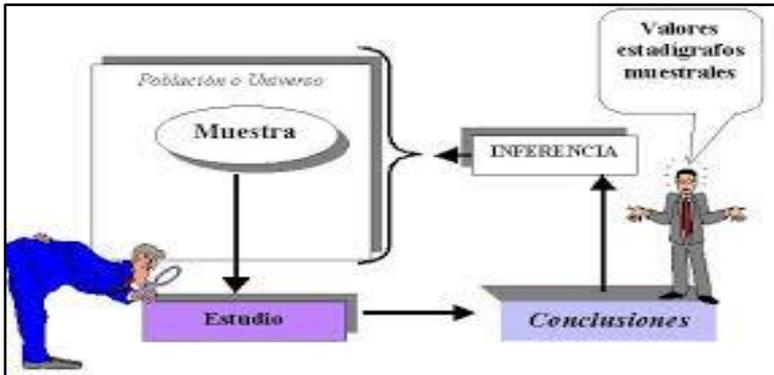


Ilustración 2: Estadística Inferencial

8. DIFERENCIA ENTRE POBLACIÓN, UNIDAD DE INVESTIGACIÓN, MUESTRA, OBSERVACIÓN

9. Población objetivo

La población objetivo se refiere a un estudio minucioso acerca de algo o alguna medición que se requiera realizar y sobre todo tomar en cuenta hacia donde se quiere dirigir, sacar conclusiones, analizar y sacar una idea general del contexto. Es decir, es un Conjunto de elementos bien definidos, sujetos a alguna medición. (Álvarez, 2007) indica, “Conjunto de elementos cosas o individuos que es objeto de interés y que se pueda estudiar en su totalidad”.

10. Unidad de investigación

Se puede decir que la unidad de investigación se relaciona con la población objetivo ya que esta efectúa una medición a los elementos que esta posee. (Hidalgo Guevara, 2009) “Da a conocer las etapas o pasos que hay que cumplir en el proceso de investigación”.

11.Muestra

Se denomina muestra a un subconjunto de “n” individuos que son parte de una investigación, pero con la particularidad de que estos individuos son tomados de la población objetivo. Según (Elmer, 2005), “Es la extracción de “n” objetos que son un caso de estudio de la población de estudio”.

12.Observación

En este caso, se resaltan todos los datos recopilados en la que se permite observar la característica de cada elemento incluidos en la muestra, (Fernández Fernández, Cordero Sánchez, Córdoba, Cordero, & Córdoba Largo, 2002) “La observación de los caracteres implica conocer la modalidad para cada elemento el valor que toma cada una de las variables o muestra de cada atributo”.

Ejemplos 1.

Tiempo (en meses) de la duración de un componente eléctrico



10	23	24	25
6	5	9	23
30	23	21	9
31	36	21	23
17	21	6	3

Muestra

13. PARÁMETRO Y ESTIMADOR

14. Parámetro

Es una cantidad Numérica calculada a partir de los elementos de una Población. Según, (Bernson 2006), “Un parámetro es una medida numérica que describe una característica de la población”.

15. Estimador o estadístico

Un estimador o también conocido como estadístico, es algo que se aproxima, no es un valor certero o concreto, por lo cual se estima una dicha información. Según (B Mode, 1982) “La palabra estimador se refiere a la función y estimación o estimado a un valor particula

r de la función. Para representar un estimador de un parámetro, colocamos un acento circunflejo sobre el símbolo del parámetro”.

Definición:

Las medidas calculadas a partir de la población, Ej. “ μ ” y “ σ ” se llaman PARÁMETROS

Las medidas calculadas a partir de las muestras, Ej. “ \bar{x} ” “S” se llaman ESTADÍSTICOS

16. TIPOS DE DATOS Y ESCALAS DE MEDICIÓN - VARIABLES

Al momento de realizar un análisis estadístico nos encontramos con dos tipos de variables: Variables cualitativas (valores no asociados a un valor numérico) y Variables cuantitativas (valores asociados a números).

17. Variables Cuantitativas

Este tipo de variables son aquellas en las que sus valores son numéricos, es decir tiene sentido hacer operaciones algebraicas, las que llevan un orden y estas tienen una clasificación discretas y continuas. Guerrero. A, Paulete. M, & Buitrago. V, (2007), indican. “Los datos recolectados cuantifican un elemento de la población o muestra”.

18. Variables Cuantitativas Discretas

Este tipo de variables son aquellas que toman valores puntuales. (Evants, 2005), indica, “Describir una distribución de la población”. Ejemplo: Edad, talla, Número de hijos, número de máquinas con problemas, cantidad de estudiantes de la carrera de Ingeniería en sistemas computacionales.

Tabla 1: Análisis de variables cuantitativas Discretas

Clase (edad)	Frecuencia Absoluta
18	2
19	15
20	13
21	1
22	6
23	2
27	1

19. Variables Cuantitativas Continuas

Es cuando, si entre dos valores, son posibles infinitos valores intermedios, es decir son aquellas variables que se pueden asumir valores de un rango o de un punto a otro. (Evans, S, & Jeffrey, 2005) “Variables continuas son aquellas que podemos medir con una precisión arbitraria que puede aumentar a medida que aumenta la precisión del instrumento de medida. Por ejemplo, la altura de una persona puede considerarse una variable continua, el tiempo de reparación de máquinas electrónicas, el tiempo de vida de un electrodoméstico.

Tabla 2: Análisis de variables cuantitativas Continuas

Tiempo de vida	Frecuencia Absoluta
[0-6]	2
[6-12]	6
[12-18]	8
[18-24]	16
[24-30]	10
[30-36]	7
[36-42]	1

20.Variables Cualitativas

Las variables cualitativas no son valores numéricos por lo tanto no se les pueden realizar operaciones algebraicas, es decir son variables que representan una cualidad o características de los elementos a tratar, sus valores no se pueden cuantificar. (Cáceres R. Á., 2007) “Las variables cualitativas consisten básicamente en una tabulación de las frecuencias y su representación gráfica mediante un diagrama de sectores o de barras”.

las variables cualitativas se clasifican en:

21.Variables Cualitativas Nominales

Las variables nominales son una clasificación de las variables cualitativas, son aquellos datos cuyos valores no se pueden ordenar. (Santo Preciado, 2010) afirma que, “Las variables nominales permiten, únicamente, definir categorías no ordenadas, cuyas posiciones pueden ser intercambiadas arbitrariamente”. Por ejemplo: el Sexo, nacionalidad, estado civil.

22.Variables Cualitativas Ordinales

Las variables Ordinales es una clasificación de las variables cualitativas, son aquellos datos en los cuales sus valores se pueden ordenar de acuerdo a un orden de prioridad, por ejemplo, el grado de satisfacción de un cierto producto, mejoría en un tratamiento clínico, satisfacción de servicio al cliente, grado de incertidumbre por la compra de un electrodoméstico. (R Johnson & Kuby, 2008) define. “Es una variable cualitativa que se presenta una posición o clasificación ordenada”

23.ANÁLISIS ESTADÍSTICO

El análisis estadístico, es el proceso de preparar los datos estadísticos con el fin de interpretarlos mediante su adecuada clasificación, su correcta asociación y su conveniente transformación en cifras relativas tales como: coeficientes, promedios, tasas, razones, entre otras, para ello se emplean operaciones matemáticas. Para determinar una adecuada clasificación de los datos se usa técnicas de distribución de frecuencias, y esto se logra mediante el uso de Tablas de Frecuencia.

24. Tablas de Frecuencia

Una tabla de frecuencias constituye un método útil y eficiente para tabular datos cualitativos y cuantitativos, y nos permite ordenar clasificar y explicar detalladamente la información anteriormente recolectada de la muestra. (Vásquez, 2012) La tabla de frecuencias está formada por los valores de una variable cuantitativa sus frecuencias correspondientes. Para poder construir la tabla de frecuencias se requieren conocer los siguientes términos. Clase, Marca de clase, Frecuencia Absoluta, Frecuencia Relativa, Frecuencia Acumulada y Frecuencia Relativa Acumulada.

Ilustración 3: Tabla de Frecuencias

Ordinal	Clase	Marca de clase	Frecuencia Absoluta	Frecuencia Acumulada	Frecuencia relativa	Frecuencia relativa acumulada
1	$[a_1, a_2)$	$(a_1+a_2)/2$	f_1	f_1	f_1/n	f_1/n
2	$[a_2, a_3)$	$(a_2+a_3)/2$	f_2	f_1+f_2	f_2/n	$(f_1+f_2)/n$
k	$[a_k, a_{k+1})$	$(a_k+a_{k+1})/2$	f_k	n	f_k/n	1



Solo para datos agrupados

25.Clase

Si tenemos datos agrupados son intervalos de igual longitud que son exhaustivos y mutuamente excluyentes.

26.Marca de clase

Para datos agrupados, es el punto medio de cada una de las clases.

27.Frecuencia Absoluta

Número de observaciones que se pueden clasificar en la clase. (Llínas, 2006), indica, “La frecuencia absoluta es el número de veces que aparece ese dato en una colección de datos”.

28.Frecuencia Relativa

Se la obtiene dividiendo que la frecuencia de la clase $i = 1, \frac{n}{\Sigma}$ (Llínas,2006), indica, “La frecuencia relativa de un dato o de una clase se encuentra dividiendo la frecuencia de dicho dato (o de la clase) entre el total de los datos”.

29.Frecuencia Acumulada

Resultado de sumar la frecuencia acumulada de la clase $(i - 1)$ con la frecuencia relativa i . (Líneas 2006) indica, “La frecuencia acumulada de cualquier dato o clase es la suma de la frecuencia de ese mismo dato o clase con las frecuencias de todos los demás o clases anteriores”.

30.Frecuencia Relativa Acumulada

Se obtiene dividiendo la frecuencia acumulada de la clase i para el total de iteraciones. (Líneas 2006) indica, “La frecuencia relativa acumulada de un dato o de una clase se obtiene dividiendo la frecuencia acumulada del dato o de una clase por el número total de datos”.

EJERCICIO 1.

El número de falta de ortografía que comete un estudiante de un dictado fue:

Análisis Estadístico Univinario

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

Variable Falta Ortográfica

n = 40

clase	Frecuencia absoluta	Frecuencia relativa	Frecuencia acumulada	Frecuencia relativamente acumulada
0	12	12/40	12	0.3
1	9	9/40	21	0.525
2	7	7/40	28	0.7
3	6	6/40	34	0.85
4	3	3/40	37	0.925

5	3	3/40	40	1
	40	Total 40/40		

31. ANÁLISIS DE DATOS CUANTITATIVOS

32. Tabla de Frecuencia por Intervalo

La distribución de frecuencias agrupadas o tabla de frecuencia con datos agrupados, se emplea si las variables toman un número grande de valores o la variable es continua. Para ello se agrupan los valores en intervalos que tengan la misma amplitud denominados clases. A cada clase se le asigna su frecuencia correspondiente

33.CONSTRUCCIÓN DE UNA TABLA DE FRECUENCIAS CUANTITATIVA

34.¿Cómo saber cuántos intervalos considerar?

Para calcular el número de intervalos y poder responder a la pregunta anterior, debemos seguir los siguientes pasos:

Calculo del rango. - Para esto debemos tener ordenados los datos.

$$R = X_{(mayor)} - X_{(menor)}$$

35.Calculo del número de intervalos. –

Para saber cuál es un número adecuado de intervalos de clase para los datos recolectados se puede aplicar una de las tres fórmulas que se presentan a continuación, esta elección depende más del problema concreto y de lo que se esté exactamente estudiando, para aquello definimos cual va hacer el n "tamaño de la muestra".

$$N_i = 1 + 3.32 \lg(n)$$

Regla de Sturges

$$N_i = 10 \lg(n)$$

Regla de Dixon y Kromal

$$N_i = 2 n^{1/2}$$

Regla de Velleman

36. Calculo del ancho del intervalo. –

Si el valor de i esta con decimales se lo redondea al valor inmediato. Ejemplo.

$$i = \frac{R}{N_i}$$

Si el valor de i esta con decimales se procede a calcular un nuevo rango.

$$\text{Nuevo}R = N_i * i$$

Si la resta entre el nuevo rango y el rango que habíamos obtenido anteriormente es:

Positivo. - Se le debe restar X al primer valor y se le debe sumar X al último valor.

Igual a 1.- Se le debe restar 0.5 al primer valor y se le debe sumar 0.5 al último valor.

Negativo. -No se pueden agrupar los datos y debemos realizar el cálculo de la tabla de frecuencia de la forma tradicional.

EJERCICIO 1.

Se tiene una muestra de 40 datos el valor máximo es 41 y el valor mínimo es 20 construya una tabla de datos agrupados y determine el número de intervalos, así como el ancho de cada intervalo.

Desarrollo.

$$n = 40$$

$$X(\text{max}) = 41$$

$$X(\text{min}) = 20$$

Calculamos el rango

1. $R = X(\max) - X(\min)$

$$R = 41 - 21$$

$$R = 21$$

Calculamos el número de intervalos, mediante la fórmula de Sturges.

2. $Ni = 1 + 3,32 * \lg(40)$

$$Ni = 6,32 \cong 6$$

Calculamos el ancho de cada intervalo

3. $i = R/Ni$

$$i = 21/6 = 3,5 \cong 4$$

Calculamos en nuevo Rango

4. $R' = Ni * i$

$$R' = 24$$

Obtenemos los intervalos

1) [19-23)

2) [23-27)

3) [27-31)

4) [31-35)

5) [35-39)

6) [39-43)

EJERCICIO 2.

A 40 estudiantes se les pide que estime el número de horas que estudiara (tanto fuera y dentro de clases) obteniendo el siguiente resultado

36	30	47	60	32	35	40	50
54	35	45	52	48	58	60	38
32	35	56	48	30	55	49	39
58	50	65	35	56	47	37	56
58	50	47	58	55	39	58	45

Elabore la tabla de frecuencia para datos agrupados, calcule la media y la varianza

1.- Ordenamos los datos

30	30	32	32	35	35	35	35	36
37	38	39	39	40	45	45	47	47
47	48	48	49	50	50	50	52	54
55	55	56	56	56	58	58	58	58
58	60	60	65					

2.- calculamos el rango, el número de intervalos, el ancho de cada intervalo

$$R = 65 - 35 \rightarrow 35$$

$$Ni = 1 + 3,32 * \lg(40) \rightarrow 6,31 \rightarrow 6$$

$$i = \frac{35}{6} = 5,83 \cong 6$$

$$R' = 6 * 6 \rightarrow 36$$

Clase	Marc a de Clase	Frecuenci a Absoluta	Frecuenci a Relativa	Frecuenci a Acumula da	Frecuenci a Relativa- Acumula da
[30- 36)	33	8	0,2	8	0,2
[36- 42)	39	6	0,15	14	0,35
[42- 48)	45	5	0,125	19	0,47
[48- 54)	51	7	0,175	26	0,65
[54- 60)	57	11	0,275	37	0,925
[60- 66)	63	3	0,075	40	1

Tabla 3: Tabla de frecuencia acumulada

Media

$$\tilde{x} \rightarrow \sum_{i=1}^n \frac{X_i}{n} = \frac{(36 + 30 + 47 + \dots + 45)}{40}$$

$$\tilde{x} \rightarrow 46,95$$

Varianza

$$S^2 \rightarrow \frac{1}{n-1} \sum_{i=1}^n (x_i - \tilde{x})^2$$

$$S^2 \rightarrow ((36 - 46,95)^2 + (30 - 46,95)^2 + \dots + (45 - 46,95)^2)$$

$$S^2 \rightarrow 9,91$$

37.GRÁFICOS ESTADÍSTICOS PARA UNA DISTRIBUCIÓN DE FRECUENCIAS CUANTITATIVA Y CUALITATIVA

Las gráficas nos permiten tener una perspectiva diferente para interpretar los datos adecuadamente. (Fernández, 2002), indica, “Nos permiten tener una perspectiva más detallada de los datos que son analizados”

38. Diagrama de Barras

Se utiliza para variables discretas con características cualitativas. Para esto se deja una separación entre las barras para indicar los valores que son posibles. Es una gráfica bidimensional en cuyo eje de las X se encuentran las clases y en el eje de las Y se encuentran la frecuencia relativa.

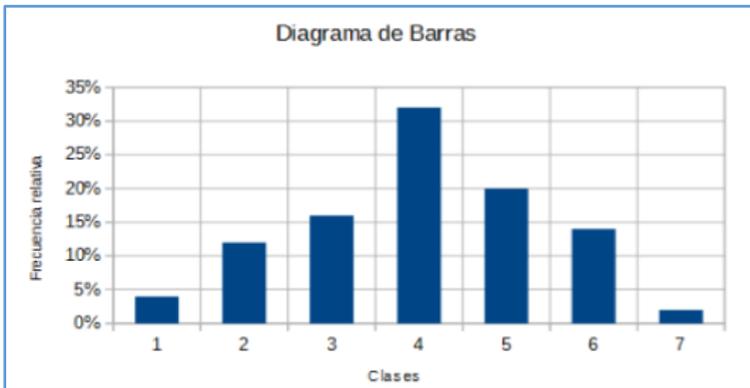


Ilustración 4: Diagrama de Barras

39. Histograma de frecuencia

Se utiliza para variables continuas no se deja separación entre las barras. Es una gráfica bidimensional en cuyo eje de las X se encuentran las clases y en el eje de las Y se encuentran la frecuencia relativa. El histograma de frecuencias es una representación visual de los datos en donde se evidencian

fundamentalmente tres características: forma, acumulación o tendencia posicional y dispersión o variabilidad.

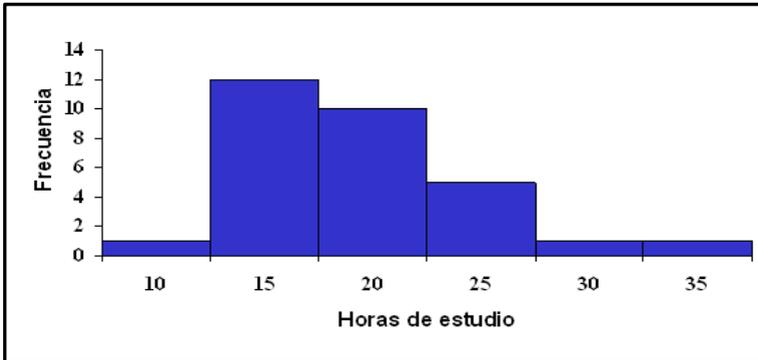


Ilustración 5: Histograma

40. Polígono de frecuencias

Gráfico bidimensional en la que el eje de las X se encuentra la marca de clase y el eje de la Y se encuentra la frecuencia relativa.

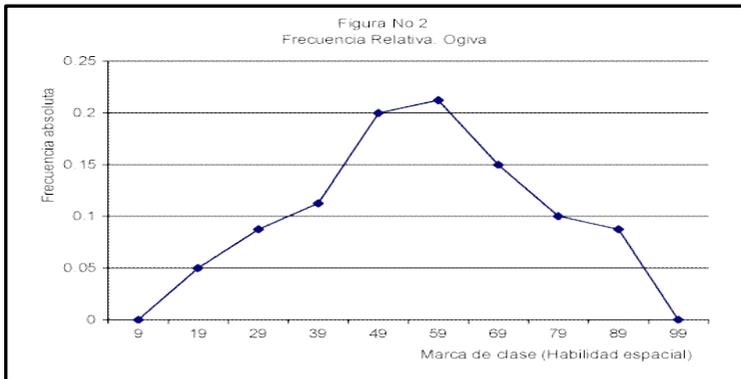


Ilustración 6: Polígono de Frecuencias

41.Ojiva

Gráfica que presenta en el eje de las X la característica cualitativa que se está investigando y en el eje de la Y se encuentra la frecuencia relativa acumulada.

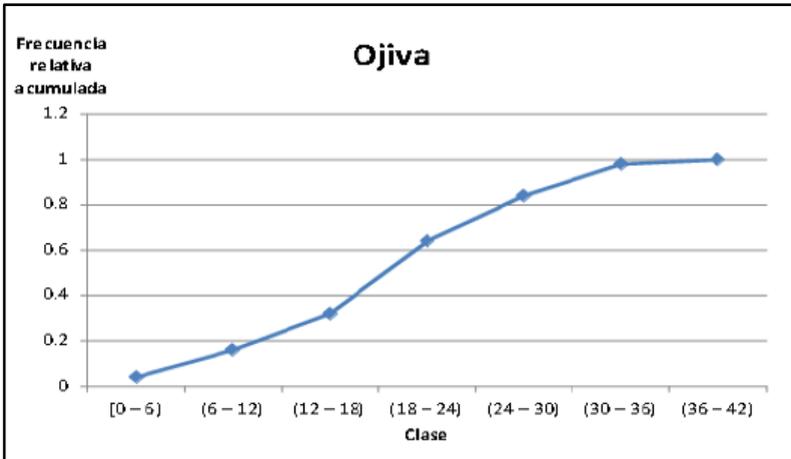


Ilustración 7: Gráfica de Ojiva

42.Diagrama de Pastel

Este tipo de gráfica se usa cuando se quiere tener una idea de la contribución de cada valor de la variable al total. Aunque es usada más para variables cualitativas, también podría usarse para variables cuantitativas discretas siempre que la variable no asumas muchos valores distintos.

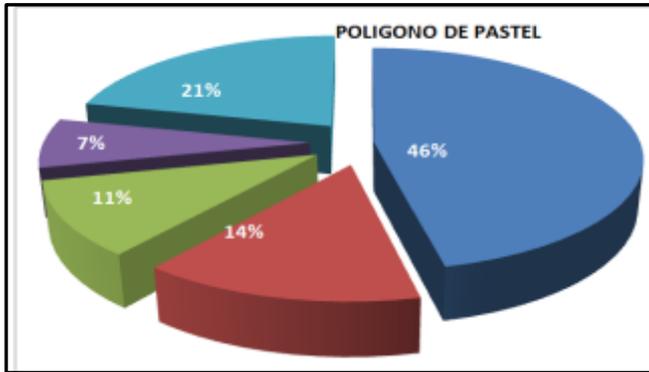


Ilustración 8: Diagrama de Pastel

Ejercicio 1.

Dada la siguiente tabla de frecuencias determine sus gráficos estadísticos (Hidograma, Polígono de frecuencia, ojiva)

Análisis Estadístico Univinario

Clase	Marca de Clase	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa - Acumulada
[18 – 21)	19,5	13	0,4333333333	13	0,4333333333
[21 – 24)	22,5	10	0,3333333333	23	0,7666666667
[24 – 27)	25,5	1	0,0333333333	24	0,8
[27 – 30)	28,5	2	0,0666666667	26	0,8666666667
[30 – 33)	31,5	3	0,1	29	0,9666666667
[33 – 36)	34,5	1	0,0333333333	30	1
		30	1		

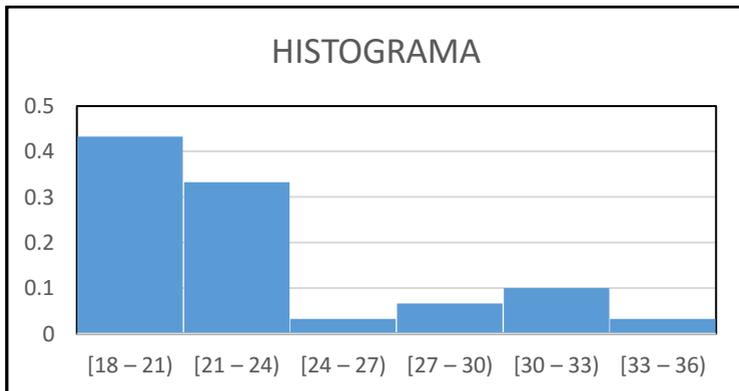


Ilustración 9: Polígono de frecuencias acumulados

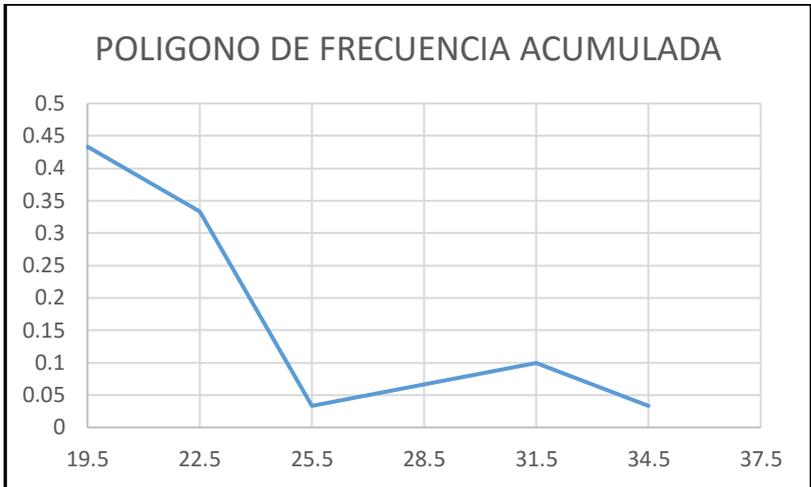


Ilustración 10: Polígono de frecuencias acumulados

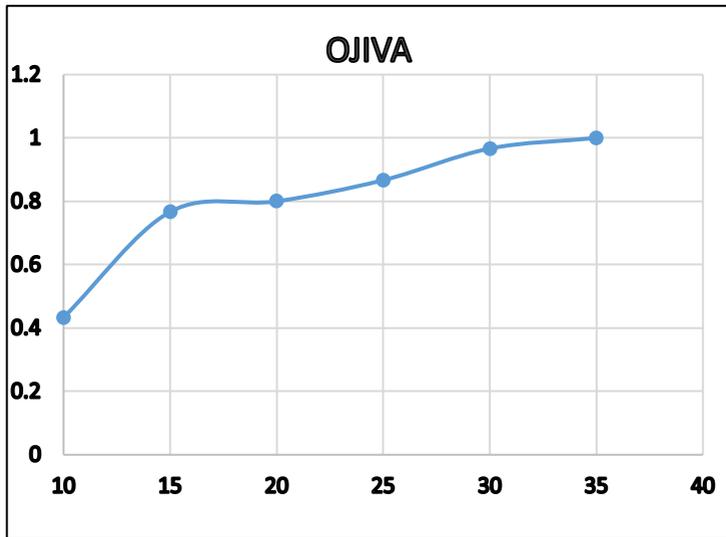


Ilustración 11: OJIVA

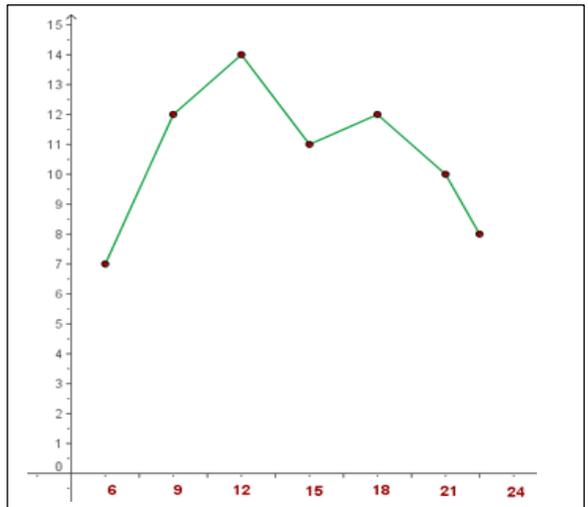
Ejercicio 2.

Análisis Estadístico Univinario

Las temperaturas en la ciudad de Quito han sufrido las siguientes variaciones, en diferentes horarios:

Represente los datos de la tabla de frecuencias, mediante un gráfico de Ojiva.

Hora	Temperatura
6	7°
9	12°
12	14°
15	11°
18	12°
21	10°
24	8°



43. Diagrama de Cajas

Gráfico que nos sirve cuando tenemos cuartiles se debe calcular los siguientes elementos:

Rango Intercuartil. - Es una medida de disposición asociada a la mediana y es la diferencia entre $Q3$ y $Q1$.

$$RI = Q3 - Q1$$

Valores Alejados. - Son valores observados que se apartan demasiado del resto de la muestra utilizamos las siguientes reglas:

1. Si el valor de X_i tomado de la muestra es menor que $Q1 - (1.5 * RI)$ entonces se dice que X_i es alejado por defecto.
2. Si el valor de X_i de la muestra es mayor que $Q3 + (1.5 * RI)$ entonces X_i es alejado por exceso.
3. Si el valor máximo de la muestra es mayor que la regla 2 se lo llama dato atípico o dato aberrante.

Los **valores atípicos** son aquellos que están más allá de los límites inferior y superior, es decir cuando más allá de 3 veces el RI en lugar del 1.5.

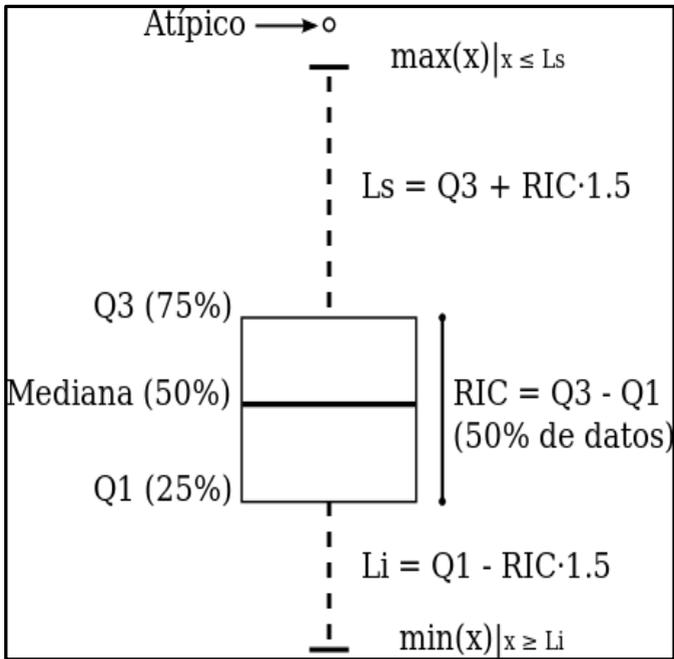


Ilustración 12: Diagrama de Cajas

Ejemplo 1:

Dada la siguiente Muestra, correspondiente a la altura media de los alumnos de una escuela particular, determinar el rango Intercuartil, el cuartil uno, dos y tres, así como el diagrama de caja y bigotes.

X1=1,65	X2=1,60	X3=1,69	X4=1,69	X5=1,81	X6=1,68	X7=1,60	X8=1,70	X9=1,70	X10=1,76
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

Ordenamos los datos de menor a mayor:

X1=1,60	X2=1,60	X3=1,65	X4=1,68	X5=1,69	X6=1,69	X7=1,70	X8=1,70	X9=1,76	X10=1,81
---------	---------	---------	---------	---------	---------	---------	---------	---------	----------

1. Calculamos los Cuartiles

$$Q1=1,63$$

$$x_{(1)}= 1,60 \quad x_{(10)}=1,81$$

$$Q2=1,68$$

$$RI=1,71-1,63= 0,08$$

$$Q3=1,71$$

2. Diagrama de cajas y bigote

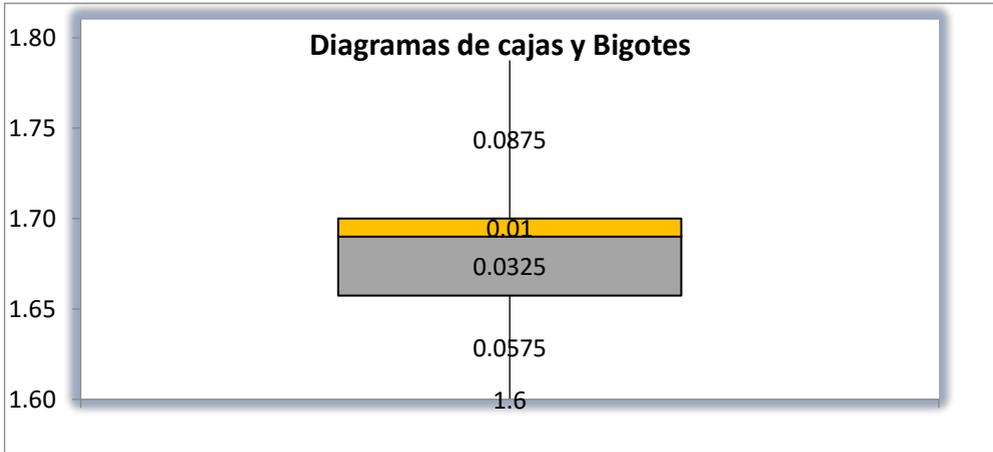


Ilustración 13: Diagrama de Cajas y Bigotes

44.MEDIDAS DE TENDENCIA CENTRAL Y DISPERSIÓN (DATOS NO AGRUPADOS)

Al describir grupos de diferentes observaciones, con frecuencia es conveniente resumir la información con un solo número, es así que las medidas de tendencia central pretenden resumir en un solo valor a un conjunto de valores. Las medidas de tendencia central más utilizadas son: media, mediana y moda.

Las medidas de dispersión pretenden evaluar en qué medida los datos difieren entre sí. De esta forma, ambos tipos de medidas usadas en conjunto permiten describir un conjunto de datos entregando información acerca de su posición y su dispersión.

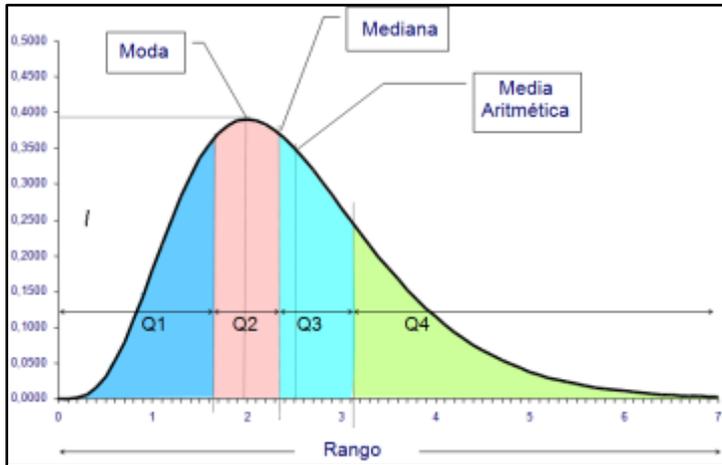


Ilustración 2: Campana de Gauss

45. Definición de estudio:

En matemática, el símbolo griego “Σ” en mayúscula se utiliza para indicar sumatoria de datos donde:

$$\sum_1^n x_i = x_1 + x_2 + x_3 + x_4 + \dots + x_n$$

Siendo “x” un valor de una medición de la variable en estudio e “i” un índice que varía de “1 a n”. El número de datos de la muestra se identifica con la letra “n”.

46. Media Muestral

La media muestral es el cálculo de la media aritmética de los datos de los cuales son objeto de estudios. Según (Luceño, 2004), “Se le llama media muestral a la media aritmética de los datos”.

$$\bar{x} = \frac{\sum_1^n x}{n}$$

47. Mediana

Es el valor que se encuentra a la mitad de los datos ordenados de una muestra, existen dos formas de calcularlo, esto dependiendo de si tenemos una cierta cantidad de datos sean estos pares o impares.

Si n es impar, la mediana es el valor que ocupa la posición $(n + 1) / 2$, una vez que los datos han sido ordenados (en orden creciente o decreciente), porque éste es el valor central. Es decir:

$$M_e = X_{\frac{(n+1)}{2}}$$

Si n es par, la mediana es la media aritmética de los dos valores centrales. Cuando n es par, los dos datos que están en el centro de la muestra ocupan las posiciones $(n / 2)$ y $(n / 2 + 1)$ Es decir:

$$M_e = X_{n/2} + X_{n/2+1}$$

48.Modas

La moda nos indica el valor que tiene mayor frecuencia absoluta en una distribución de datos. (George, 2003), “La moda muestra hacia que valor tienden los a agruparse. En conjunto relativamente pequeños, puede que no exista un par de observaciones cuyo valor sea el mismo”.

48.1 Otras medidas de tendencia central.

48.1.1 La Media Geométrica.

La media geométrica se define como $\bar{x}_g = \sqrt[n]{x_1 x_2 x_3 \dots x_n}$, por ejemplo, la media geométrica de los valores “4, 5, 4, 6” es

$$\bar{x}_g = \sqrt[4]{(4)(5)(4)(6)} = 4,68$$

48.1.2 La Media Cuadrática.

Se construye a partir de suma de los cuadrados de un conjunto de

valores. Su forma de cálculo es $\bar{x}_c = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}{n}}$, si tomamos los valores anteriores la media cuadrática tiene el

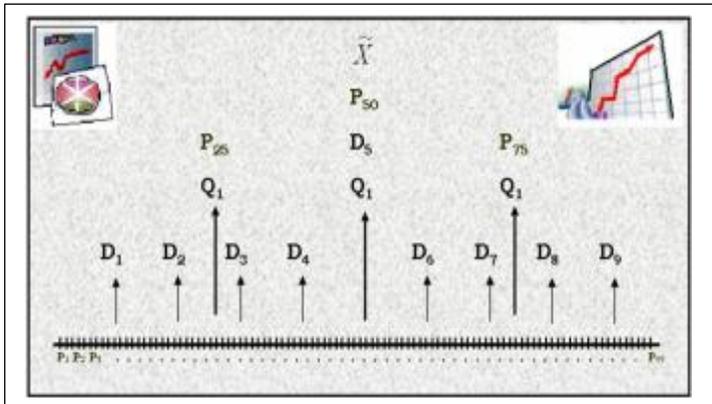
siguiente valor
$$\bar{x}_c = \sqrt{\frac{4^2 + 5^2 + 4^2 + 6^2}{4}} = 4.81$$

49.MEDIDAS DE POSICIONAMIENTO

Las medidas que describen un valor típico en un grupo de observaciones suelen llamarse medidas de posicionamiento. Es

importante tener en cuenta que estas medidas se aplican a grupos más bien que a individuos.

Entre las medidas de posicionamiento tenemos:



Deciles (D_i)

Percentiles (P_i)

Cuartiles (Q_i)

Ilustración 15: Percentiles,

Deciles, Cuartiles

50. CUANTILES

Los Cuantiles se relaciona con el concepto de los cuartiles, también intervienen los deciles, percentiles, es una distribución de partes iguales, también se los conoce como centiles.

(Manteiga, M. T. G., 2012) “Los Cuantiles, también llamados centiles, son valores que dividen la distribución en partes de igual frecuencia. Los más importantes son la mediana, los cuantiles, los deciles y los percentiles”.

$$P_{(i)} = x \left(\frac{(n + 1) * i}{100} \right)$$

$$I = 1 \text{-----} 99$$

$$X_{(i,a)} = x_{(i)} + 0. a(x_{(i+1)} - x_{(i-1)})$$

$$I = \textit{Parte entera}$$

$$A = \textit{Parte decimal}$$

51.PERCENTILES, ORDEN I

Los percentiles son las distribuciones en centésimas al cual se lo representa con $i\%$ que es la que me indica el valor a calcular, para poder realizar estos debe haber por lo menos 100 datos y si no existen se necesita hacer aproximaciones, en otros casos se conoce a los percentiles como los 99 valores que divide la serie, el percentil equivale al 1% del total de observaciones. Según (Alvarado Valencia, J. A., & Obagi Araújo, J. J., 2008) El percentil Divide la distribución de la muestra en centésimas partes, es decir, existen 99 percentiles o centiles, también como lo indica (Martin A. 2004), “Cuando la muestra ordenada se divide en 100 partes iguales se obtienen los percentiles”.

$$\text{Posición de } P_i = \frac{i * (n - 1)}{100} + 1$$

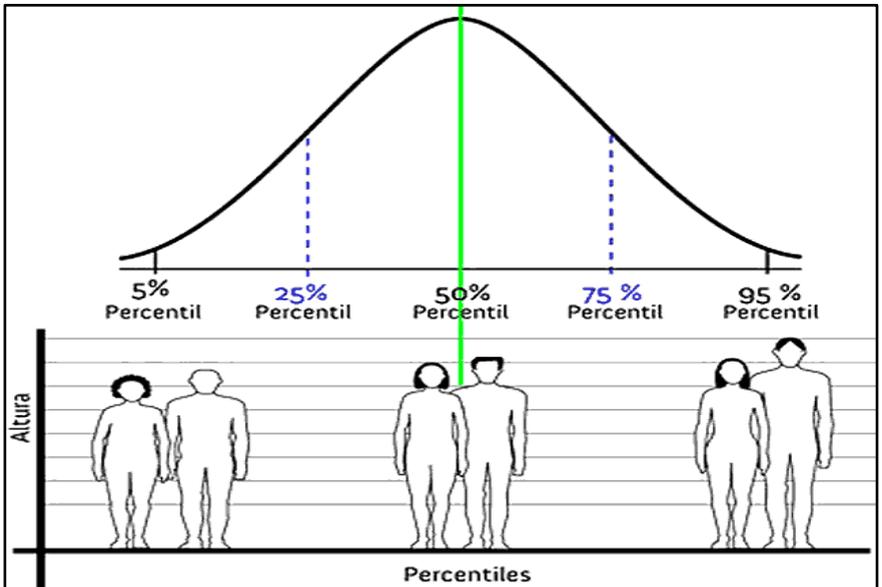


Ilustración 3: División de la muestra en 100 partes

iguales

52.DECILES

Los deciles se divide a la distribución en grupos de 10 partes iguales, cada parte de los deciles contiene el 10% de los datos acumulados. (Muñoz, D. R., 2005) "Son los valores de la variable que dividen a la distribución en partes iguales, cada una de las cuales engloba el 10% de los datos.

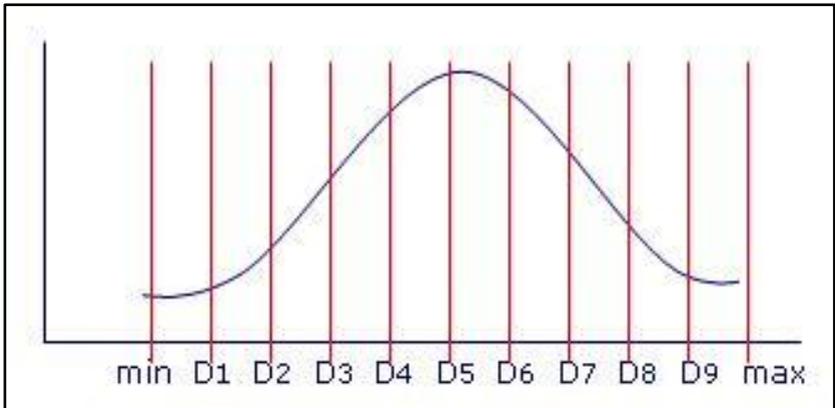


Ilustración 4: División de la muestra en 10 partes iguales

53. CUARTILES

Los cuartiles, divide los elementos de la muestra en 4 grupos con frecuencias similares. (Bernson 2006), indica “Los cuartiles dividen a un conjunto de datos en 4 partes iguales”.

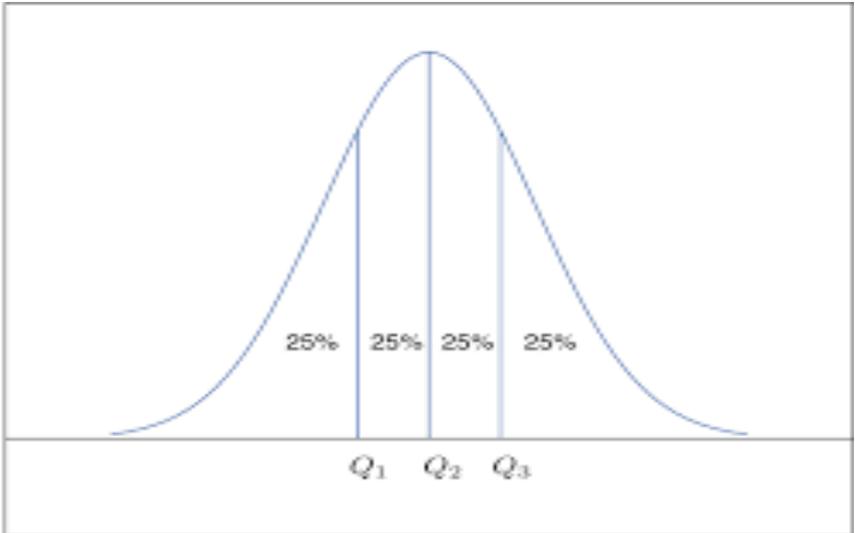


Ilustración 5: División de la muestra en 4 partes iguales

Para una mejor comprensión, se muestra en la tabla #, una comparación de cada uno de los estadísticos de posición (percentiles, deciles, centiles) con respecto a los cuartiles (Q_i).

TABLA DE ASOCIACIONES			
	PERCENTILES	DECILES	CENTILES
Q1 = 25%	25	2.5	0.25
Q2 = 50%	50	5	0.5
Q3 = 75%	75	7.5	0.75
Q4 = 100%	100	10	1

Para hallar un estadístico de posición aplicaremos la siguiente formula:

$$X(i.a) = X_{(i)} + 0.a (X_{(i+1)} - X_{(i)})$$

$$P_i = \left(\frac{(n+1) i}{100} \right)$$

$i = \text{Posición real del elemento}$
 $a = \text{Parte decimal}$

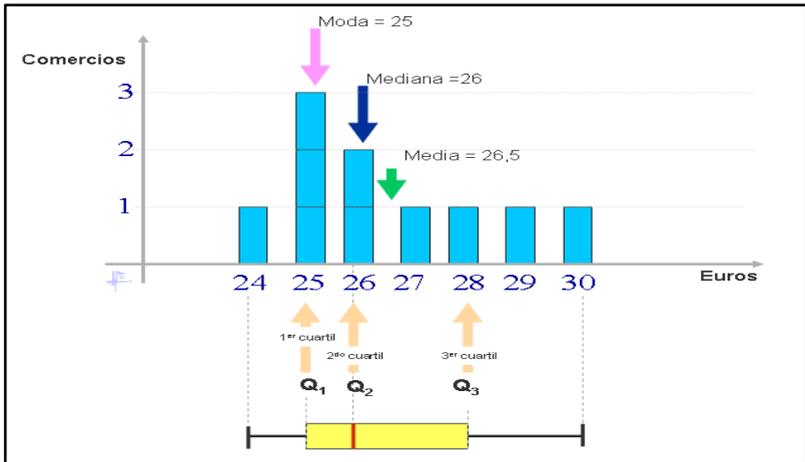


Ilustración 19: Diagrama de cajas y Barras

Ejemplo 1:

Se dan las edades de un grupo de estudiantes de la carrera de Lenguas:

21, 20, 22, 20, 20, 25, 23

a. Se procede a ordenar los datos

Tabla 4: Edades.

Edades	
$X_{(1)}$	20
$X_{(2)}$	20
$X_{(3)}$	20
$X_{(4)}$	21
$X_{(5)}$	22
$X_{(6)}$	23
$X_{(7)}$	25

b. Mediante la fórmula de los percentiles encuentre la posición de los cuartiles.

$$Q1 \text{ -- } 25\% = 20$$

$$Q2 \text{ -- } 50\% = 21$$

$$Q3 \text{ -- } 75\% = 23$$

$$P_{(i)} = x \left(\frac{(n + 1) * i}{100} \right)$$

$$i = 1 \text{ --- } 99$$

$$X_{(i,a)} = x_{(i)} + 0.a(X_{(i+1)} - X_{(i-1)})$$

$i = \text{Parte entera}$

$A = \text{Parte decimal}$

Ejercicio 1:

Tabla de datos desordenada

a1a	
X ₍₁₎	1.65
X ₍₂₎	1.60
X ₍₃₎	1.64
X ₍₄₎	1.69
X ₍₅₎	1.81
X ₍₆₎	1.68
X ₍₇₎	1.60
X ₍₈₎	1.70
X ₍₉₎	1.70
X ₍₁₀₎	1.76

Tabla de datos ordenada

Talla	
X ₁	1.60
X ₂	1.60
X ₃	1.64
X ₄	1.65
X ₅	1.68
X ₆	1.69
X ₇	1.70
X ₈	1.70
X ₉	1.70
X ₁₀	1.81

Tabla 3: Tablas de Datos.

$$Q1 = 25\% = P(25)$$

$$P_{(25)} = x \left(\frac{(11) * 25}{100} \right) =$$

$$X_{(2,75)} = x_{(2)} + 0.75(X_{(3)} - X_{(2)})$$

$$X_{(2,75)} = 1.60 + 0.75(1.64 - 1.60)$$

$$X_{(2,75)} = 1.60 + 0.75(0.04)$$

$$Q2 = 50\% = P(50)$$

$$P_{(50)} = x \left(\frac{(11) * 50}{100} \right) =$$

$$P_{(50)} = x(5.5)$$

$$X_{(5,5)} = x_{(5)} + 0.5(X_{(6)} - X_{(5)})$$

$$X_{(5,5)} = 1.68 + 0.5(1.69 - 1.68)$$

$$X_{(5,5)} = 1.68 + 0.05(0.01)$$

$$Q3 = 75\% = P(75)$$

$$P_{(75)} = x \left(\frac{(11) * 75}{100} \right) =$$

$$P_{(75)} = x(8.25)$$

$$X_{(8,25)} = x_{(8)} + 0.25(X_{(9)} - X_{(8)})$$

$$X_{(8,25)} = 1.70 + 0.25(1.76 - 1.70)$$

$$X_{(8,25)} = 1.70 + 0.25(0.26)$$

Ejercicios Propuestos

1. Encuentre los cuartiles Q1, Q2, Q3, P70 de los siguientes datos:

Datos	Datos ordenados
4 7 2 9 3	2 3 4 7 9 → $n = 5$

2. Encuentre Q1, Q2, Q3, P65, P70 para los siguientes datos:

Datos	Datos ordenados
36 25 37 24	20 23 24 24
39 20 36 45	24 25 29 31
31 31 39 24	31 33 34 36
29 23 41 40	36 37 39 39
33 24 34 40	40 40 41 45

54.MEDIDAS DE DISPERSIÓN O VARIABILIDAD

5.1 El Rango.

El Rango, Recorrido o Amplitud de un conjunto de mediciones, “es la diferencia entre el valor mayor y el valor menor”, indica el número necesario y mínimo de unidades, en la escala de medición, para incluir los valores mínimo y máximo. Es la medida de dispersión más fácil de calcular, pero también es la menos estable al estar fuertemente influenciada por valores extremos atípicos.

$$R = X_{m\acute{a}ximo} - X_{m\acute{i}nimo}$$

5.1.1 Varianza Muestral

La varianza muestral nos indica la desviación de los datos con respecto a la media, como sabemos la variación de una constante es cero porque nunca cambiara su valor.

$$S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$$

5.2 Desviación Estándar

La desviación estándar o desviación típica es la raíz cuadrada de la media de los cuadrados de las puntuaciones de desviación (varianza), se representa por σ .

$$S = \sqrt{\sum_1^n \frac{(x_i - \bar{x})^2}{n - 1}}$$

5.3 Coeficiente de Variación (C.V)

El coeficiente de variación permite comparar las dispersiones de dos distribuciones distintas, siempre y cuando ambas sean positivas, se obtiene entre la relación de la desviación estándar y su media.

$$C.V = \frac{\text{Desviación Estandar}}{\text{Media}} * 100\%$$

$$C.V = \frac{S}{\bar{x}} * 100\%$$

Donde:

C.V = Coeficiente de variación.

S= desviación estándar de la muestra.

\bar{x} = media aritmética de la muestra

Ejemplo 1.

En la tabla siguiente se tiene las notas del primer parcial de un examen de algebra lineal, correspondiente a 10 estudiantes elegidos al azar de un paralelo, de la carrera de ingeniería en Sistemas computacionales de la universidad de Guayaquil, la escala de calificación del examen es de 1 a 100, determine la media la desviación estándar y la varianza

Estudiante	“Variable Nota = x_i ”	Valor de x_i
Luis	X_1	62
Alberto	X_2	68
Juan	X_3	92
Pedro	X_4	88
Roberto	X_5	55
María	X_6	79
Raquel	X_7	89
Luisa	X_8	92
Rosa	X_9	67
Diana	X_{10}	69
	$\sum_1^{10} x_i =$	761.

Solución:

El valor correspondiente a la Media de notas de los estudiantes es:

$$\bar{x} = \frac{\sum_1^{10} x_i}{n} = \frac{62+68+92+88+55+79+89+92+67+69}{10}$$

$$\bar{x} = 761/10$$

$$\bar{x} = 76.1$$

El valor de la Desviación Estándar está dado por:

$$S = \sqrt{\sum_1^n \frac{(x_i - \bar{x})^2}{n - 1}}$$

$$\sqrt{\frac{((55 - 76.1)^2 + (62 - 76.1)^2 + (67 - 76.1)^2 + (68 - 76.1)^2 + (69 - 76.1)^2 + (79 - 76.1)^2 + (88 - 76.1)^2 + (89 - 76.1)^2 + (92 - 76.1)^2 + (92 - 76.1)^2)}{9}}$$

$$S =$$

$$S = 13.6$$

El valor de la Varianza, está dado por:

$$S^2 = 184.96$$

Ejemplo 2.

Del ejercicio correspondiente a las notas del primer parcial de un examen de algebra lineal, correspondiente a 10 estudiantes elegidos al azar de un paralelo de la carrera de ingeniería en Sistemas computacionales de la universidad de Guayaquil, determine la mediana y la moda.

Solución: Para poder determinar la mediana, el primer paso es ordenar los datos de menor a mayor, a continuación, pasamos a ordenar los datos en la siguiente tabla:

Datos Ordenados

Estudiante	"Datos ordenados"	Valor de x_i
Roberto	1	55
Luis	2	62
Rosa	3	67
Alberto	4	68
Diana	5	69
María	6	79
Pedro	7	88
Raquel	8	89
Juan	9	92
Luisa	10	92

Como "n" es par, la mediana es igual a la mitad entre la medición con rango " $n / 2$ " y la medición con rango " $(n/2) + 1$ ", donde " $n / 2$ " = 5 y " $(n / 2) + 1$ " = 6.

El dato 5 vale 69 y el dato 6=79, entonces "la mediana" es igual a $69 + 79 / 2 = 74$

En este ejemplo para poder determinar el valor de la Moda, nos damos cuenta que el valor que mayormente se repite dentro la muestra es $M_o = 92$

Ejemplo 3.

Andrés, es un estudiante de la universidad de Guayaquil, tiene las siguientes calificaciones en las 10 asignaturas que

recibe en su carrera: 8, 7, 10, 9, 8, 7, 8, 10, 9 y 10. Alberto, un amigo de Andrés, que estudia en la Universidad Católica, y tiene las siguientes calificaciones: 8, 9, 8, 7, 8, 9, 10, 7, 8 y 10. Se desea determinar a partir de la información proporcionada por el departamento de asesoría académica. ¿Cuál estudiante tiene menor variabilidad en sus calificaciones?

Solución:

Como se está tomando en cuenta todas las asignaturas, se debe calcular el coeficiente de variación poblacional.

La media para Andrés es:

$$\bar{x} = \sum_1^{10} x_i / n = \frac{8+7+10+9+8+7+8+10+9+10}{10}$$

$$\bar{x} = \sum_1^{10} x_i / n = 8,6$$

La media para Alberto es:

$$\bar{x} = \sum_1^{10} x_i / n = \frac{8+9+8+7+8+9+10+7+8+10}{10}$$

$$\bar{x} = \sum_1^{10} x_i / n = 8,4$$

La desviación estándar para Andrés es:

$$S = \sqrt{\sum_1^n \frac{(x_i - \bar{x})^2}{n - 1}}$$

$$S = \sqrt{\frac{12,4}{9}} = 1.37$$

La desviación estándar para Alberto es:

$$S = \sqrt{\frac{10,4}{9}} = 1.08$$

El coeficiente de variación para Andrés es:

$$C.V = \frac{S}{\bar{x}} * 100\%$$

$$C.V = \frac{1.37}{8.6} * 100\% = 15.81$$

El coeficiente de variación para Alberto es:

$$C.V = \frac{1.08}{8.4} * 100\% = 12.86$$

Por lo tanto, el estudiante que tiene menor variabilidad en sus calificaciones es Andrés

6 MEDIDAS DE FORMA

6.1 Coeficiente de Asimetría (g1)

El coeficiente de asimetría g_1 se define como: Grado en que una distribución de puntos está concentrada en un extremo o en el otro

$$g_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{s^3}$$

Cuando el coeficiente de asimetría (g_1), es igual a cero se dice que es simétrico, en este caso tenemos que la media, la mediana y la moda son iguales.

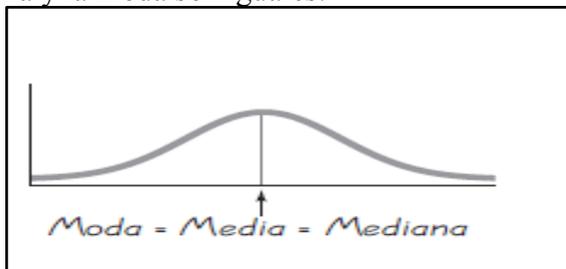


Ilustración 20: Distribución Simétrica

Cuando el coeficiente de asimetría (g_1) es mayor a cero, entonces decimos que la distribución Es asimétrica hacia la

derecha (tiene sesgo positivo o es sesgada a la derecha). Es decir, la media es mayor a la mediana.

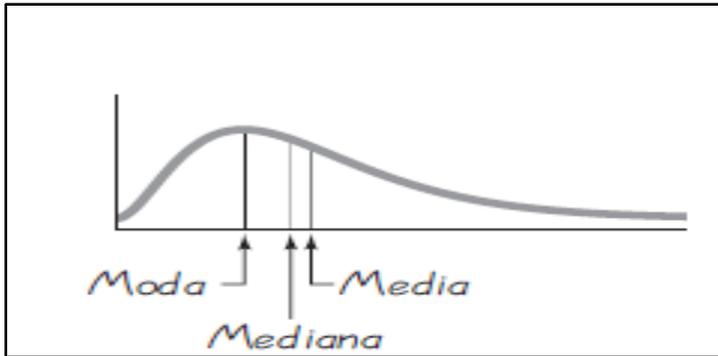


Ilustración 21: Distribución Asimétrica hacia la Derecha

Cuando el coeficiente de asimetría (g_1) es menor a cero, entonces decimos que la distribución es asimétrica hacia la izquierda (tiene sesgo negativo o sesgado a la izquierda). Es decir, podemos decir que la media es menor a la mediana.

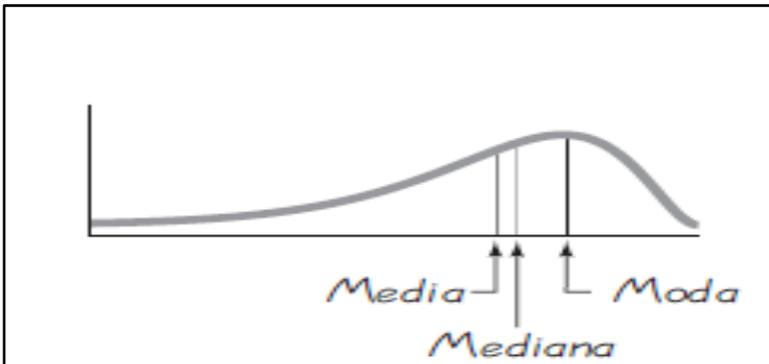
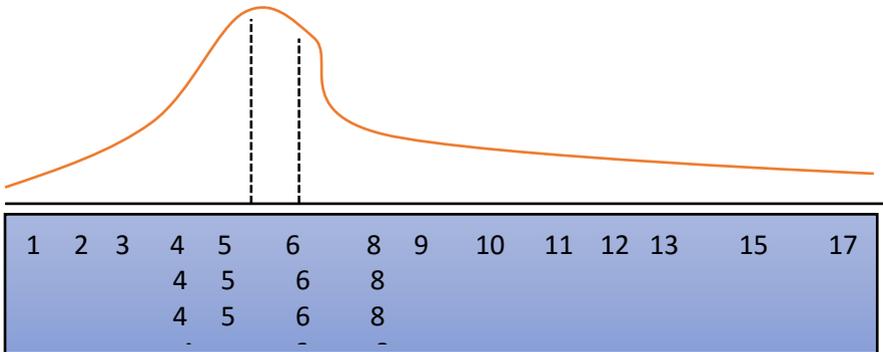


Ilustración 22: Distribución asimétrica hacia la Izquierda

Ejemplo 1: Dada la siguiente muestra, determine el tipo de asimetría, teniendo en consideración, los datos correspondientes a la diferencia que existe entre la media y la mediana

1,2,3,4,4,4,4,5,5,5,6,6,6,6,8,8,8,8, 9,10,11, 12,13,15,17

$\tilde{X} = 6 < \bar{x} = 7.2$. Para este caso tenemos que la Mediana es menor que la media, es decir tenemos una distribución asimétrica hacia la derecha, esto implica que los datos se encuentran mayormente agrupados hacia la izquierda de la campana de Gauss.



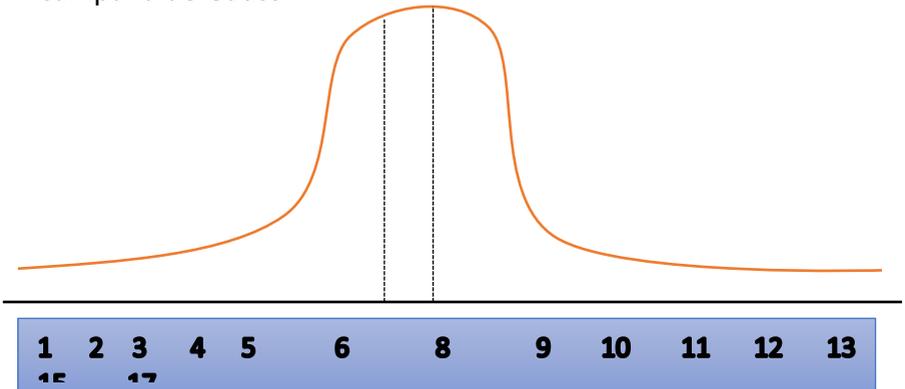
$$\tilde{X} = 6 < \bar{x} = 7.2.$$

Ilustración 23: Asimétrica hacia la derecha

Ejemplo 2: Dada la siguiente muestra, determine el tipo de asimetría, teniendo en consideración, los datos correspondientes a la diferencia que existe entre la media y la mediana

1,2,3,4,5,6,9,9,12,12,12,15,17

$\bar{x} = 8.23 < \tilde{X} = 9$. Para este caso tenemos que la Media es menor que la mediana, es decir tenemos una distribución asimétrica hacia la izquierda, esto implica que los datos se encuentran mayormente agrupados hacia la derecha de la campana de Gauss.



$$\bar{x} = 8.23 < \tilde{X} = 9.$$

Ilustración 24: Distribución asimétrica

6.2 Coeficiente de Curtosis o Aplastamiento (g2)

La Curtosis es el estadístico de forma que nos permite saber si la distribución de los datos de una muestra esta agrupada o no alrededor de la media, según la observación de su Campana de Gauss. Con el coeficiente de Curtosis, podemos identificar si existe una gran concentración de valores alrededor de la media. El coeficiente de Curtosis o aplastamiento g2 se define como:

$$g_2 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{s^4} - 3$$

para ello la Curtosis g2 se la mide a través de tres formas diferentes:

Si $g_2 = 0$ (distribución Mesocúrtica), el grado de aplastamiento de los datos será similar al de una campana de Gauss.

Si $g_2 > 0$ (distribución Leptocúrtica), los datos representan un menor aplastamiento que la campana de Gauss.

Si $g_2 < 0$ (distribución Platicúrtica), los datos aparecen más aplastados que la campana de Gauss.

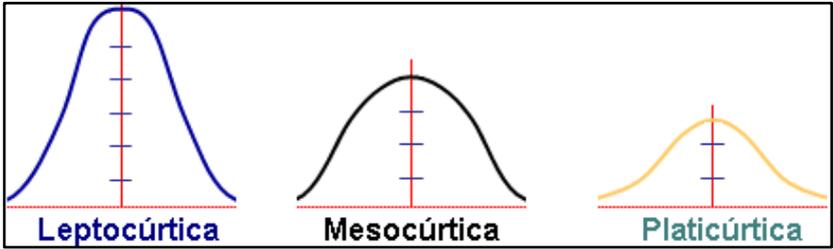


Ilustración 25: Formas de Curtosis

7 DISEÑO MUESTRAL

7.1 Definición de Muestreo

El muestreo se define como un procedimiento mediante el cual se obtienen datos que posteriormente serán analizados, para determinar el grado de exactitud de una población objetivo con la finalidad de poder inferir sobre los resultados de dicha población. (Macassi & Cristina, 1997) ""Consiste en un conjunto de reglas, procedimientos y criterios mediante los cuales se selecciona un conjunto de elementos de una población que representan lo que sucede en toda esa población".

7.2 Tamaño de la Muestra

7.2.1 Población Finita

Es un conjunto con un número limitado de elementos. La fórmula para hallar el tamaño de la muestra cuando se conoce la población es la siguiente:

$$n = \frac{NZ_{\alpha}^2\sigma^2}{(N-1)E^2+Z_{\alpha}^2\sigma^2} \text{ O } n = \frac{NZ_{\alpha}^2PQ}{(N-1)E^2+Z_{\alpha}^2PQ}$$

Como podemos observar, tenemos dos fórmulas, de las cuales solo para este libro utilizaremos la formula dos, que corresponde a los valores de la varianza proporcional. Esto se da cuando no tenemos mucha información que nos ayude a poder calcularla.

A continuación, presentamos la correspondiente interpretación que contiene cada una de las variables, en donde:

N = Tamaño de la población.

E = Error. $E=5\% - 10\%$

Z = Nivel de Confianza. $\rightarrow Z=95\%; \alpha=0.05; Z_{\alpha}=1.96$

Varianza

$\sigma^2 \rightarrow$ Cuando conocemos \bar{X} y S^2
 $PQ \rightarrow$ Cuando no conocemos \bar{X} y S^2

$$\bar{X} = \sum \frac{X_i}{n}$$

$$\sigma^2 = S^2 = \frac{1}{n-1} * (X_1 - \bar{X})^2$$

P = éxito $\rightarrow 0.5$

Q = fracaso $\rightarrow 0.5$

Ejemplo: Encuentre el tamaño de la muestra sabiendo que:

$$N = 43,700$$

$$Z_{\alpha} = 95\% \rightarrow 0.95$$

$$E = 5 \rightarrow 0.05$$

$$n = \frac{N * Z_{\alpha}^2 * P * Q}{(N - 1) * E^2 + Z_{\alpha}^2 * P * Q}$$

$$n = \frac{(43,700) * (3.84) * (0.5) * (0.5)}{(43,700 - 1) * (0.00025) + (3.84) * (0.5) * (0.5)}$$

$$n = \frac{41,952}{109.2475 + 0.96}$$

$$n = \frac{41,952}{110.2075}$$

$$n = 380.66$$

7.2.2 Población Infinita

Es aquella que tiene un número muy extenso de elementos, la fórmula para hallar el tamaño de la muestra cuando la población es infinita, es la siguiente:

$$n = \frac{Z_{\alpha}^2 * p * q}{d^2}$$

En donde:

q = probabilidad de fracaso

Z = Nivel de Confianza → *Z* = 95 % ; $\alpha = 0.05$; $Z_{\alpha} = 1.96$

d = Margen de error

p = probabilidad de éxito o proporción esperada

Ejemplo 1:

¿A cuántas familias tendríamos que estudiar para conocer la preferencia del mercado en cuanto a las marcas de Shampoo para bebe, si se desconoce la población total?

Seguridad = 95%

Precisión = 3%

Proporción Esperada = Asumamos que puede ser próxima al 5%; si no tuviésemos idea de dicha proporción utilizaríamos el valor de $p = 0.5$ (50%) que maximiza el tamaño Muestral.

Entonces:

$$Z_{\alpha}^2 = 1,962 \text{ (ya que la seguridad es del 95\%)}$$

$$p = 0,05 \text{ (5\%)}$$

$$q = 1 - p$$

$$d = 0,03 \text{ (3\%)}$$

$$n = \frac{1,96^2 * 0,05 * 0,95}{0,03^2} = 203$$

Se requeriría encuestar a no menos de 203 familias para poder tener una seguridad del 95%.

7.3 Tipos de Muestreos: Muestreos no Probabilísticos

El muestreo no probabilístico se caracteriza por que la muestra se obtiene a partir de criterios establecidos, este tipo de muestreo es muy poco utilizado debido a que la muestra extraída no resulta representativa sin embargo se lo utiliza cuando el presupuesto de una investigación es reducido ya que en comparación con el muestreo probabilístico su costo es considerablemente mínimo. (Torres, 2007) “Este tipo de muestreo se utiliza cuando el probabilístico resulta muy costoso, teniendo presente que no sirve para hacer generalizaciones puesto que no existe certeza de que la muestra extraída tenga representatividad, puesto que no todos los elementos de la población tienen la misma probabilidad de ser seleccionados”.

7.3.1 Muestreo por Cuotas

Consiste en seleccionar una muestra con la peculiaridad que previamente se divide la población en estratos de cada estrato se selecciona una muestra, la cual debe de contar con determinadas características (o también llamadas cuotas) para dicho estudio, en este tipo de muestreo el investigador debe de tener una sólida base de conocimientos de cada estrato y el criterio de selección lo realiza dependiendo del estudio a realizar (edad, sexo, economía, etc.). En primera instancia se puede llegar a confundir con el Muestreo Probabilístico Estratificado, pero la diferencia es que la selección en este tipo de muestreo no se lo realiza de manera aleatoria. (Rabolini, 2009) “Consiste en formar estratos de la población sobre la base de ciertas características y en procurar que estén representadas en proporciones semejantes a las que existen en la población “.

1. La población se divide en k estratos como por ejemplo (la edad, el sexo, colores favoritos, etc.)

$$N = N_1 + N_2 + N_3 + \dots + N_k$$

2. El investigador elige las cuotas, siendo la suma total el tamaño de la muestra

$$n = n_1 + n_2 + n_3 + \dots + n_k$$

3. Se seleccionan los elementos en cada estrato de manera no probabilística.



Ilustración 26: Representación Gráfica del muestreo por Cuotas

7.3.2 Muestreo de Opinión o Intencional

Se determina la muestra por criterios propios del investigador, no existe criterios de formación y tampoco criterios de matemáticas. Se recomienda que este tipo de muestreo sea utilizado en estudios de nivel micro.

Un claro ejemplo de este tipo de muestreo se lo vive en las aulas escolares cuando el docente decide formar grupos de estudio y le provee toda la libertad para que los estudiantes escojan a su grupo de estudio. Cada estudiante realiza un muestreo debido a determinados criterios (siempre trabaja con determinado sujeto, se sabe desenvolver, tiene una excelente oratoria, su promedio es excelente, etc.).

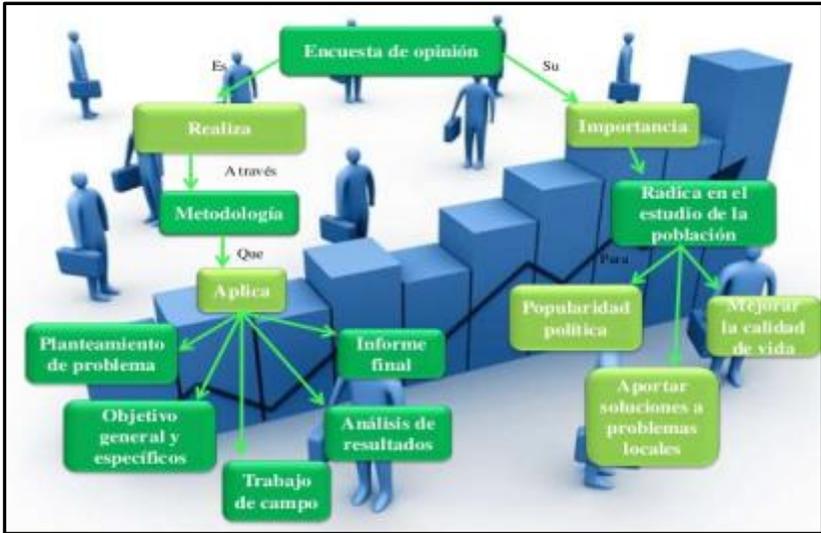


Ilustración 27: Representación Gráfica del muestreo por Opinión

7.3.3 Muestreo Casual o Incidental

Este tipo de muestreo se caracteriza porque no existe un plan de formación de la muestra, simplemente el investigador define el sector, se dirige con su grupo de trabajo y allí realiza su estudio a los individuos que de manera inoportuna se acerquen, dichos individuos pasan a formar parte de la muestra. Es considerado uno de los muestreos menos aplicables debido a que no se obtiene el resultado deseado. (Padilla & Gil, 2007) “Para realizar un estudio determinado, el investigador elige, por ejemplo, los alumnos de un centro que se encuentra próximo a su lugar de trabajo. Este tipo de muestreo se considera el más alejado de la posibilidad de generalizar a la población los resultados obtenidos”.



Ilustración 28: Representación Gráfica del muestreo Casual

7.3.4 Muestreo Bola de Nieve

El muestreo Bola de Nieve o también denominado Muestreo por redes tiene la peculiaridad de empezar por una muestra y que dicha muestra le permita de manera directa o indirecta encontrar a otra muestra, se crea una red hasta que se pueda determinar un número de muestra considerable. Este tipo de muestreo se utiliza mucho cuando dificulta mucho el acceso a la población. (Rabolini, 2009) “Se basa en redes sociales, en las amistades. Cuando se encontró el primer representante, éste puede conducir a otro, y éste a un tercero, y así sucesivamente hasta conseguir una muestra suficiente “.



Ilustración 29: Representación Gráfica del muestreo bola de nieve

7.3.5 Muestreo Discrecional

También denominado muestreo por juicio, en este tipo de muestreo el investigador es el que selecciona la muestra debido a su criterio y experiencia en estudios similares realizados a la misma población o una población distinta, la aplicación de este muestreo se la realiza cuando la población es reducida. Las desventajas que presenta este muestreo se refleja en la fiabilidad y el sesgo que acompaña.

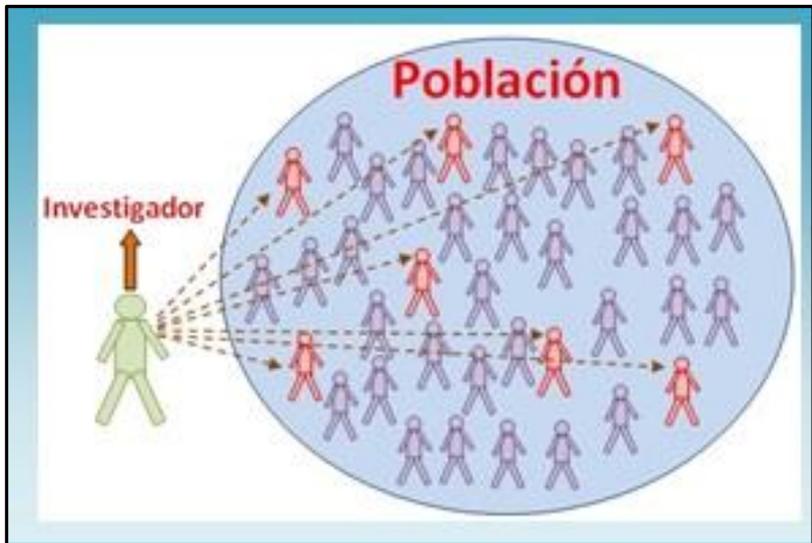


Ilustración 30: Representación Gráfica del muestreo Discrecional

Al momento de realizar una investigación de campo es necesario realizar un estudio de todos los datos necesarios obtenidos a partir de una muestra, en ocasiones dicho estudio resulta muy complejo por el gran número de datos a manipular, como respuesta a esto surge el muestreo probabilístico el cual consiste en obtener la muestra, aplicar uno o más métodos para que cualquier individuo de la muestra tenga igual probabilidad de ser elegido. (Lagares & Albandoz, 2001) “Es aquel en el que cada muestra tiene la misma probabilidad de ser elegida”.

7.3.6 Muestreo Aleatorio Simple

Cada elemento de la población es elegido de manera aleatoria, dichos elementos tienen la misma probabilidad de ser elegidos. Para lograr establecer una muestra se realiza una determinación de que elemento será el estudiado. (Webster, 2001) “Garantiza que cada muestra de un tamaño dado tenga la misma probabilidad de ser seleccionada”.

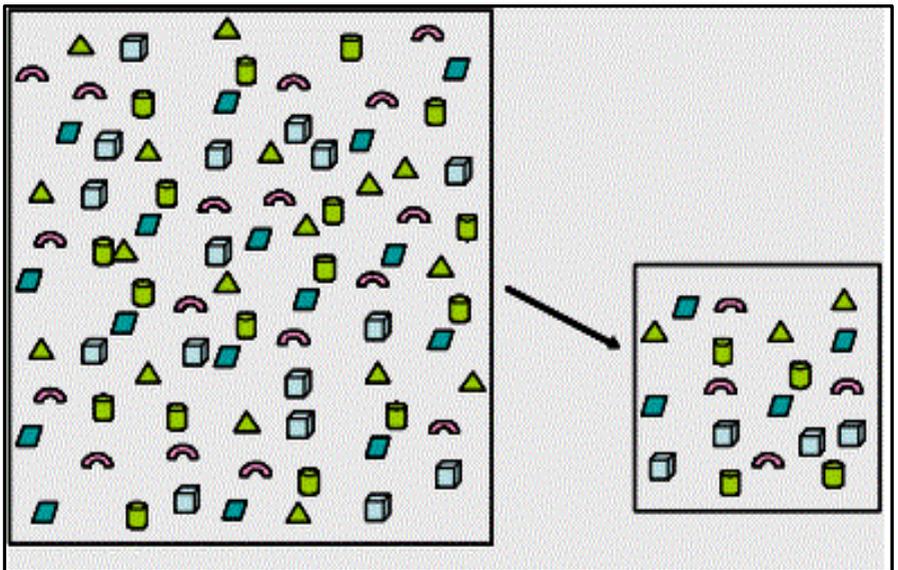


Ilustración 31: Representación Gráfica del muestreo aleatorio Simple

Ejemplo 2:

En una encuesta realizada a los estudiantes de la Universidad de Guayaquil se obtuvieron los siguientes datos:

N°	Nombre	Edad
1	Maria	20
2	Pedro	32
3	Juan	19
4	Francisco	17
5	Angie	24
6	Lisbeth	22
7	Estefania	23
8	Rafael	19
9	Raul	18
10	Julieta	28
11	Guisela	27
12	Martha	30

Asumiendo que el tamaño de la muestra es de 4 elementos aplicamos muestreo Aleatorio simple, a partir de números aleatorios generados desde Excel con la fórmula:

=ALEATORIO.ENTRE(i
nicio; fin)

N°	Aleatorio	Nombre	Edad
1	1	Maria	20
2	8	Rafael	19
3	10	Julieta	28
4	12	Martha	30

7.3.7 Muestreo Estratificado

Su principal característica es la de subdividir la población total en subpoblaciones, dichas subpoblaciones se denominan estratos y la sumatoria de los estratos es el tamaño total de la población; dicho de otra manera, $N=N_1+N_2+N_3+ \dots + N_k$. Cada estrato está compuesto por una muestra y la sumatoria de esta muestra tiene que ser igual al tamaño de la muestra previamente calculado $n=n_1+n_2+n_3+ \dots + n_k$. (Webster, 2001). “Se toma una muestra estratificada forzando las proporciones de la muestra de cada extracto para que este forme el patrón poblacional”.

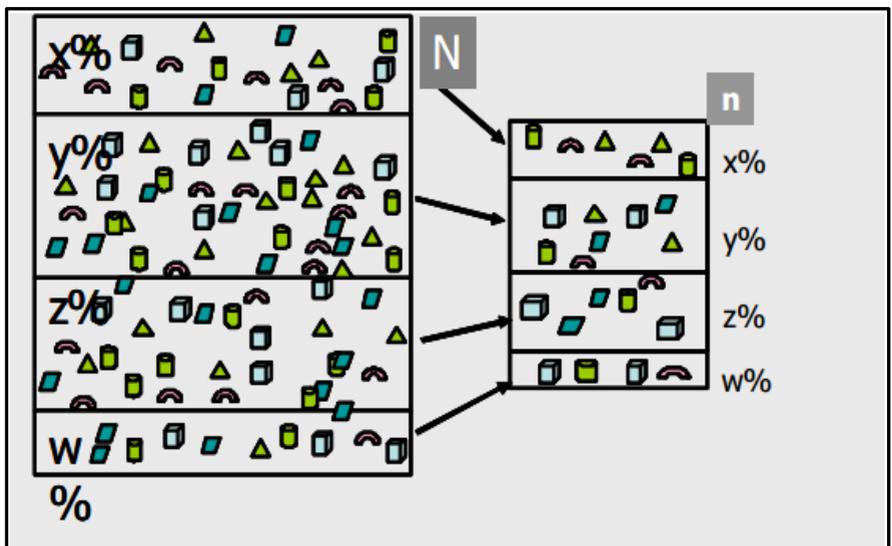


Ilustración 32: Representación Gráfica del muestreo aleatorio estratificado

La fórmula para determinar los estratos es la siguiente:

$$ni = n * \frac{Ni}{N}$$

Dónde:

ni = Número de Estratos

n = Tamaño de la Muestra

Ni = Tamaño de Población en cada estrato

N = Tamaño de la Población

Ejemplo 3:

En una fábrica que consta de 600 trabajadores queremos tomar una muestra de 20. Sabemos que hay 200 trabajadores en la sección A, 150 en la sección B, 150 en la sección C y 100 en la sección D.

En el ejercicio tenemos nuestro tamaño de la población $N=600$, al igual que el tamaño de la muestra $n=20$; cómo podemos observar la población está dividida en 4 extracto $N1=200$, $N2=150$, $N3=150$, $N4 =100$. Entonces para escoger mi extracto de interés aplico la fórmula del muestreo estratificado para cada una de subpoblaciones o extracto de población.

Estrato A= $N1$

Estrato B= $N2$

Estrato C= N3

Estrato D= N4

Para $N_1=200$;

$$n1 = 20 * \frac{200}{600} = 6,67 \cong 7$$

Para $N_2=150$;

$$n2 = 20 * \frac{150}{600} = 5$$

Para $N_3=150$;

$$n3 = 20 * \frac{150}{600} = 5$$

Para $N_4=100$

$$n4 = 20 * \frac{100}{600} = 3,33 \cong 3$$

De los siguientes resultados se deduce que, para un objeto de estudio se tomaran 7 elementos del Estrato A, 5 elementos del Estrato B, 5 elementos del Estrato C y 3 elementos del Estrato D.

7.3.8 Muestreo por Conglomerados

Existen casos en que comúnmente la población se encuentra de manera natural agrupada por conglomerados, para ejemplarizar esto nos podemos fijar en una Urbanización la misma se encuentra dividida en ciudadelas, manzanas y villas si quisiéramos realizar un estudio de satisfacción con respecto al servicio de seguridad, podríamos elegir cualquier conglomerado al azar y estudiar todos los elementos para dicho estudio. (Casal & Mateu,

2003) “Se divide la población en varios grupos de características parecidas entre ellos y luego se analizan completamente algunos de los grupos, descartando los demás”.



Ilustración 33: Representación Gráfica del muestreo por conglomerados

7.3.9 Muestreo Sistemático

Para la realización de un muestreo sistemático se realiza una división de la población sobre la muestra, el resultado de esta operación se lo conoce como un coeficiente de elevación el cual nos servirá como punto de partida para utilizar posteriormente un Muestreo Aleatorio Simple y así seguir determinando los elementos de la muestra hasta su tope. (Torres, 2007) “Consiste en dividir la población en n estratos, compuestos por las primeras K unidades, las segundas k unidades y así sucesivamente. Otra definición del muestreo sistemático es la que fue dada por (Casal & Mateu, 2003) “Se elige el primer individuo al azar y el resto viene condicionado por aquél”.

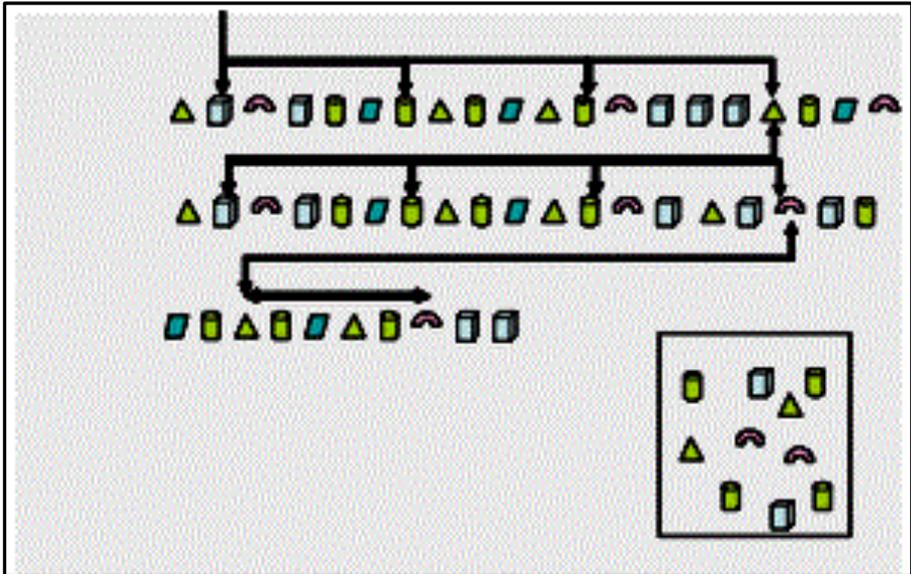


Ilustración 34: Representación Gráfica del muestreo aleatorio sistemático

La fórmula para determinar el coeficiente de Elevación es la siguiente:

$$k = \frac{N}{n}$$

En donde:

k = Coeficiente de Elevación

n = Tamaño de la Muestra

N = Tamaño de la población

Ejemplo:

A partir de una población de 100 estudiantes de la Universidad Laica se desea estimar una muestra de 20 alumnos.

$$k = \frac{100}{20} = 5$$

Se elige un número aleatorio entre 1 y 5. Asumiremos que dicho número es el 4. Los elementos para el estudio serían los siguientes:

2,7,12,17,22,27,32,37,42,47,52,57,62,67,72,77,82,87,92,97

7.4Elaboración del Cuestionario

El cuestionario se define como una herramienta utilizada por los investigadores para la recolección de datos y posterior estudio de los mismos, como punto de partida tenemos que definir qué tipo de investigación se desea realizar y dependiendo de aquello elaborar nuestro cuestionario. Esta herramienta resulta muy eficaz al momento de medir una magnitud de un determinado Fenómeno social, el origen, los sucesores e impacto de dicho fenómeno. A continuación, detallamos algunos ejemplos en los que se puede realizar un cuestionario:

- ✓ ¿Cuántas personas creen que el método de estudio vía web es más eficiente que el método de estudio de manera presencial? Interesa saber la magnitud
- ✓ ¿Cuál es el método de comunicación que utilizan las personas actualmente: ¿Correo electrónico, video

llamadas, foros o chats? En este caso nos interesa saber cómo se origina un suceso.

- ✓ ¿Cuáles son las razones para que las personas no usen Internet? Aquí nos interesa saber la sucesión de un fenómeno.

(Martinez F. , 2002) Señala 4 pasos a considerar en la elaboración de un cuestionario:

1. Describir la Información que se necesita.
2. Redactar las preguntas y escoger el tipo de Preguntas.
3. Redactar un texto introductorio y las instrucciones.
4. Diseñar el aspecto formal del cuestionario.

Existen preguntas abiertas y preguntas cerradas el uso de cada una depende de la investigación que se va a realizar.

7.4.1 Preguntas Abiertas

Las preguntas abiertas no suelen tener alternativas, constan de un espacio donde la persona que es objeto de estudio podrá redactar con sus propias palabras acerca de la cuestión previamente planteada, este tipo de preguntas se utilizan cuando el investigador no puede prever las posibles respuestas que el encuestado puede tener; estas preguntas resultan más fáciles de plantear, pero el análisis de cada pregunta resulta muy complejo.

Ejemplo 5:

¿Por qué asiste a Psicoterapia?

¿Qué opina de los dibujos animados que actualmente se transmiten por Televisión?

7.4.2 Preguntas Cerradas

Este tipo de preguntas contienen categorías previamente delimitadas, el tiempo que se tarda el investigador en elaborar este tipo de preguntas con sus posibles respuestas en considerablemente alto sin embargo el análisis de las mismas se da en un menor tiempo en comparación con las preguntas abiertas ya que son datos que se pueden cuantificar. La principal desventaja de utilizar solo preguntas cerradas en cuestionario es que limita al encuestado a responder acerca de un determinado tema y en ocasiones las categorías planteadas no son acordes a las respuestas del encuestado. (Hernandez, Fernandez, & Baptista, 2003) “Cada tipo de preguntas tiene sus ventajas y desventajas, la elección de uno u otro tipo depende del grado en que se puedan anticipar las posibles respuestas, el tiempo del que se dispone para codificar y si se quiere una respuesta más precisa o profundizar en alguna cuestión”.

Este tipo de preguntas se clasifican en:

- ✓ Dicotómicas y excluyentes
- ✓ Varias alternativas de respuesta y excluyentes
- ✓ No excluyentes

Ejemplo 6:

7.4.3 Preguntas Dicotómicas y excluyentes

¿Estudia usted actualmente?

- Si
- No

7.4.4 Preguntas de Respuesta Múltiple y Excluyente

¿Cuál es el puesto que ocupa en su empresa?

- Director General
- Gerente
- Subgerente
- Coordinador
- Jefe de Área
- Supervisor
- Empleado
- Otro

7.4.5 Preguntas No excluyentes

Esta familia tiene: (puede escoger más de una opción)

- Radio
- Televisión
- Teléfono
- Automóvil
- Ninguna de las Anteriores

7.5 Elaboración del Informe

Un informe es una herramienta donde se describe de manera oral o escrita todo el trabajo realizado sin la necesidad de especificar cálculos estadísticos ya que el mismo puede ir dirigido a distintas áreas que no tengan correlación con la parte estadística. Según (Cassany, 2007) un informe debería de contener los siguientes campos:

7.5.1 Título:

Debe de dar una idea clara del tema el mismo debe de ser concreto y atractivo para el lector.

7.5.2 Resumen:

Se deben de explicar todas las situaciones y conclusiones más relevantes acontecidas durante el trabajo.

7.5.3 Introducción:

Por qué se realizó el estudio, describo el objetivo y alcance de todo el estudio.

7.5.4 Métodos:

Describe la manera en cómo se llegan a los resultados, cabe recalcar que el informe debe de permitir que otras personas puedan repetir el experimento sobre la base del mismo. Se sintetiza Origen, Herramientas y Diseño.

7.5.5 Resultados:

Establece lo que se ha encontrado, aquí se detallan de manera gráfica los resultados obtenidos.

7.5.6 Análisis:

Interpreta los datos obtenidos del estudio y realiza todas las recomendaciones posibles.

7.5.7 Conclusiones



**PARTE II:
ESTUDIOS DE
CASOS**

INTRODUCCIÓN

La estadística es una disciplina importante, aporta a la formación y al desarrollo de las competencias de un alumno con el perfil de Ingeniero en Sistemas computacionales. A partir de ello y como parte de las estrategias que la cátedra permanentemente revisa y mejora, y teniendo muy presente el uso de herramientas Ofimáticas actuales por parte de los alumnos, se consideró como parte de la gestión académica dentro del aula de clases, proponer proyectos a los estudiantes, en la que su base principal era el análisis estadístico descriptivo, aplicados a casos de estudio de índole académico y pedagógico, para aquellos estudiantes que cursan la asignatura Probabilidad y Estadística en la carrera.

En sus aspectos metodológicos se crean temas relacionados a la índole pedagógica, se formula una serie de preguntas que son plasmadas en un formulario de preguntas para encuestas; para luego crear en Excel una “Base de datos” que presenta ciertas variables de una población real, a partir de la cual se deberán

aplicar las herramientas estadísticas para obtener conclusiones o realizar inferencias, generando los informes de análisis correspondientes. Permite el uso adecuado de las herramientas estadísticas con interpretaciones apropiadas al contexto del problema y el conocimiento.

El Libro, “Análisis Estadístico Univariado, casos de estudio”, que se pone a disposición del estudiante, presenta un conjunto de sugerencias, para orientarlo en el desarrollo del trabajo práctico. En este Libro se sugiere la utilización de bibliografía, las consultas a los docentes en horas de clases prácticas, las consultas a tutorías en los laboratorios informáticos.

Con este libro, “Análisis Estadístico Univariado, casos de estudio”, también se hace entrega al alumno, un instructivo didáctico y estructurado para la presentación de trabajos relacionados con estadística, indicando la estructura exigida para su presentación impresa y digital.

8 “ANÁLISIS ESTADÍSTICO PARA DETERMINAR LA PREVALENCIA DEL PLAGIO ACADÉMICO EN TRABAJOS DE INVESTIGACIÓN CIENTÍFICA”

En la actualidad el plagio académico es una práctica muy frecuente en los estudiantes universitarios. Las nuevas tecnologías y el acceso a internet les facilita dicha información haciendo que alumno tenga la facilidad del “copia y pega” en trabajos de investigación y se olvide de citar dichos conceptos, poniendo en duda la credibilidad de los autores y de las instituciones que los respaldan.

Se considera como plagio al copiar un proyecto, o una parte del mismo y presentarlo como propia. Utilizar gráficos, frases, textos, sin citar su procedencia; así como descargar trabajos de internet y solo cambiarle el nombre de los autores eso es plagiar; incluso

copiar nuestros propios trabajos y presentarlos como innovadores también es plagio, pero en este caso sería auto plagio.

El tema del plagio académico, cada vez se profundiza más, por la facilidad que los estudiantes tienen al acceso del internet. Es importante que se analice esta problemática, ya que cada vez es más frecuente, se debería inculcar la honestidad académica y crear un hábito de no copiar, en los estudiantes de enseñanza superior.

Está investigación tiene como propósito, mediante un análisis estadístico descriptivo, dar a conocer a los estudiantes una estadística real, del porcentaje elevado que existe en las diferentes universidades, tomadas como objeto de estudio, de cómo sus estudiantes han cometido plagio académico mediante el uso libre que les facilita el internet.

8.1 DISEÑO DE LA INVESTIGACIÓN

El proyecto planteado sirve como parte de una solución viable que permite establecer estadísticamente la prevalencia del plagio académico, dicho estudio da solución a un problema de carácter práctico, que otorgan beneficios en diferentes áreas, nuestro campo de estudio es aplicado a la Universidad Laica Vicente Rocafuerte y la Universidad de Guayaquil

8.2 POBLACION Y MUESTRA

8.2.1 POBLACIÓN OBJETIVO

Los individuos que van a formar parte de nuestra encuesta para nuestro caso de estudio, los alumnos de la universidad Laica Vicente Rocafuerte, universidad de Guayaquil:

Tabla 1: Cuadro Poblacional

Población	N°
Universidad Laica Vicente Rocafuerte	20.000

Tabla 2: Cuadro Poblacional de las Facultades de la Universidad Laica Vicente Rocafuerte

Población	N°
Facultad de Ciencias Administrativas	10.000
Facultad de Ciencias de la Educación	100
Total	10.100

Tabla 3: Cuadro Poblacional de las carreras de la ULVR

Población	N°
Mercadotecnia	700
Inglés	53
Total	753

Tabla 4: Cuadro Poblacional de las Facultades de la Universidad de Guayaquil

Población	N°
Facultad de Ciencias Administrativas	10.000
Total	10.000

Tabla 5: Cuadro Poblacional de las carreras de la Universidad de Guayaquil

Población	N°
Ingeniería en Marketing y Negociación Comercial	250
Ingeniería Comercial	696
Total	946

8.2.2 TAMAÑO DE LA MUESTRA

P= Probabilidad de éxito (0.50)

Q = Probabilidad de fracaso (0.50)

N = Tamaño de la población ()

E = Error de estimación (5%)

Z= Valor de z, 1.96 para $\alpha = 0.05$ y 2.58 para $\alpha = 0.01$

n = Tamaño de la muestra (350)

8.2.3 CÁLCULO DEL TAMAÑO DE LA MUESTRA

$$n = \frac{z^2 * PQ * N}{(N - 1) * E^2 + z^2 * PQ}$$

$$n = \frac{(1.96)^2 * 0.25 * 10000}{(100000 - 1) * (0.05)^2 + (1.96)^2 * 0.25}$$

$$n = \frac{96.040}{(99999) * 0.0025 + 0.9604}$$

8.3 INSTRUMENTOS DE RECOLECCION DE DATOS

Una de las técnicas de campo que se va a utilizar, es la técnica de la encuesta la cual nos permitirá obtener información necesaria acerca del comportamiento del fenómeno que está sujeto a nuestro estudio, a continuación, presentamos la encuesta que se realizó a los alumnos de la universidad Laica Vicente Rocafuerte y la Universidad de Guayaquil, quienes fueron nuestra población objetivo.

8.3.1 ENCUESTA

A.- BÁSICO DE PREGUNTAS

1.- Género del entrevistado: **M**___ **F**___

2.- Edad del entrevistado (en años): _____

3.- Indique el nombre de la Institución de Educación Superior a la que Pertenece:

4.- Facultad a la que Pertenece: _____

5.- Carrera a la que pertenece: _____

6.- En que semestre/año usted se encuentra Actualmente (si es Semestre, considere el que tenga más del 50% de las asignaturas tomadas):

7.- Indique Cuantas asignaturas está tomando en este semestre/año (de 1 a 7):

8.- Indique si en las asignaturas que está cursando, le envían trabajo de investigación Académica o científica.

Sí _____ No _____

9.- En caso de que su respuesta sea “Sí”, Indique en cual (es) de las asignaturas que usted está cursando le envían más trabajos de investigación (Máximo 4):

1. _____ 2. _____ 3. _____ 4. _____

10. Indique si de las asignaturas que está cursando. ¿Algún docente tiene grado académico de PhD?

Sí _____ cuantos () No _____

11.- Indique el lugar donde usted realiza sus trabajos de investigación Académica o científica

Hogar ___ Universidad ___ Trabajo: ___ Cyber ___ Otros (indique) _____

12.- Con qué frecuencia usted realiza trabajos de investigación académica o científica.

Siempre ___ Algunas Veces ___ A menudo ___ Rara vez ___
Nunca ___

13.- Cuando usted realiza un trabajo de investigación académica, ¿Qué tipos de fuentes utiliza?

Internet __ Libros de textos __ Artículos de revistas __ Biografías __
Otros _____ (indique)

14.- Dentro del Internet ¿Qué fuente de información, considera usted confiable para realizar sus trabajos de investigación académica?

Rincón Del Vago __ Monografias.Com __ SlideShare.net __
BuenasTareas.com __
Otros _____ (indique).

15.- Dentro del Internet ¿Cuáles de las Bases de Datos Científicas conoce o utiliza para hacer sus trabajos de Investigación Académica?

Redalyc __ Scielo __ Dialnet __ Google Académico __ Ninguna __

16.- Conoce Usted, si la Institución de Educación Superior en la que estudia, posee Base de Datos Científicas (Bibliotecas Virtuales) para sus trabajos de investigación.

SI __ NO __

17.- En caso de que su respuesta a la pregunta anterior sea "Si", indique. ¿Qué Base de Datos científica utiliza con mayor frecuencia? (puede marcar más de una)

Taylor & Francis __ Springer __ Gale __ ProQuest __ Otras
_____ (indique)

18.- ¿Por qué motivos usted realiza un trabajo de investigación académica?

Cumplir con la tarea __ Conocimiento Propio __ Hacer un
buen trabajo __
Obtener Una Calificación __ Otros
_____ (indique)

19.- ¿Cuántas Horas semanales dedica usted, para hacer su trabajo de investigación?

[0 – 1] __ Hr.

[2 – 3] __ Hr.

[4 – 5] __ Hr.

[6 – 7] __ Hr.

Mayor a 8 __ Hr.

20.- Cuando Obtiene la información que buscaba para su trabajo de investigación académica, ¿usted se toma el tiempo necesario para leerla y analizarla detenidamente?

Si __ No __

21.- Cuando finaliza su trabajo de investigación ¿Cree usted que adquirió los conocimientos adecuados acerca del tema que investigó?

Total Acuerdo __ Parcial Acuerdo __ Indiferente __ Parcial Desacuerdo
__ Total Desacuerdo __

**B.- CARACTERÍSTICAS CREENCIAS O COMPORTAMIENTOS DE
LOS ALUMNOS AL INVESTIGAR**

22.- *Considera Usted que la falta de tiempo debido a asuntos personales o de trabajo, es un factor que le impide realizar una mejor investigación a tal punto que decide construir sus trabajos tomando fragmentos de diversos sitios del internet.*

Total Acuerdo __ **Parcial Acuerdo** __ **Indiferente** __ **Parcial Desacuerdo** __
Total Desacuerdo __

23.- *¿Considera Usted que el docente de la asignatura que le envía una investigación, le hace un seguimiento de su trabajo a través de alguna herramienta que le permita detectar el porcentaje de coincidencia bibliográfica con otros trabajos académicos?*

Total Acuerdo __ **Parcial Acuerdo** __ **Indiferente** __ **Parcial Desacuerdo** __
Total Desacuerdo __

24.- *¿Considera Usted que sus trabajos de investigación no aportan nada a su formación profesional?*

Total Acuerdo __ **Parcial Acuerdo** __ **Indiferente** __ **Parcial Desacuerdo** __
Total Desacuerdo __

25.- *¿Considera Usted que todo lo que está en internet se puede copiar?*

Total Acuerdo __ **Parcial Acuerdo** __ **Indiferente** __ **Parcial Desacuerdo** __
Total Desacuerdo __

26.- *¿Considera Usted que es importante realizar citas bibliográficas, en un trabajo de investigación?*

Total Acuerdo __ **Parcial Acuerdo** __ **Indiferente** __ **Parcial Desacuerdo** __
Total Desacuerdo __

27.- *¿Cuánto tiempo le dedica usted a la lectura semanalmente?*

[0 – 2] __ **Horas/semana.**

[3 – 5] __ **Horas/semana.**

[6 – 8] __ **Horas/semana.**

Más de _____ **Horas/semana.**

No leo _____

28.- *Cuanto tiempo le dedica a la investigación de los trabajos enviados por su profesor.*

[0 – 2] __ **Horas**

[3 – 5] __ **Horas**

[6 – 8] __ **Horas**

Más de _____ **Horas**

No investigo _____

29.- *¿Sabía usted que copiar un trabajo de otra persona, se considera plagio académico?*

Total Acuerdo __ **Parcial Acuerdo** __ **Indiferente** __ **Parcial Desacuerdo** __
Total Desacuerdo __

30.- *¿Usted cree que la causa de copiar un trabajo académico por parte de los alumnos es debido a la facilidad que ofrece el internet para encontrar información?*

Total Acuerdo __ **Parcial Acuerdo** __ **Indiferente** __ **Parcial Desacuerdo** __
Total Desacuerdo __

31.- ¿Usted cree que la causa de copiar un trabajo académico por parte de los alumnos es debido a que todo lo que se encuentra en internet se puede utilizar, por ser un acceso público?

Total Acuerdo __ Parcial Acuerdo __ Indiferente __ Parcial Desacuerdo __ Total Desacuerdo __

32.- ¿Usted cree que la causa de copiar un trabajo académico por parte de los alumnos, es debido a que **NO** sabe cómo realizar trabajos de investigación académica?

Total Acuerdo __ Parcial Acuerdo __ Indiferente __ Parcial Desacuerdo __ Total Desacuerdo __

8.3.2 TABULACIONES DE LA ENCUESTA

Tabla 6: Tabulaciones de la encuesta

Pregunta	Opciones	Resp	Porc j	Total
Género del entrevistado	M	107	31%	350
	F	243	69%	
Institución de Educación Superior a la que Pertenece:	Universidad Laica Vicente Rocafuerte	189	54%	350
	Universidad de Guayaquil	161	46%	
Facultad a la que Pertenece	Facultad De Ciencias De La Educación Mención En Ingles	53	15%	350
	Facultad De Ciencias Administrativas	297	85%	
Carrera a la que pertenece:	Ingeniería Comercial	69	20%	350
	Ingeniería En Marketing Y Negociación Comercial	92	26%	
	Ingles	53	15%	

Análisis Estadístico Univinario

	Mercadotecnia	136	39%	
<i>En que semestre/año usted se encuentra Actualmente</i>	Primer semestre	16	5%	350
	Segundo Semestre	68	19%	
	Tercer Semestre	17	5%	
	Cuarto Semestre	40	11%	
	Quinto Semestre	56	16%	
	Séptimo semestre	113	32%	
	Octavo Semestre	22	3%	
	Noveno Semestre	18	5%	
<i>Indique Cuantas asignaturas está tomando en este semestre/año (de 1 a 7)</i>	1	27	7%	350
	2	9	3%	
	3	49	14%	
	4	118	34%	
	5	20	6%	
	6	40	11%	
	7	87	25%	
<i>Le envían trabajo de investigación Académica o científica.</i>	Si	14	4%	350
	No	336	96%	
<i>¿Algún docente tiene grado académico de PhD?</i>	Si	128	37%	350
	No	221	63%	
<i>Donde usted realiza sus trabajos de investigación Académica o científica</i>	Hogar	244	70%	350
	Universidad	140	40%	
	Trabajo	48	14%	
	Cyber	79	23%	
	Otros	3	1%	
<i>Con qué frecuencia usted realiza trabajos de investigación académica o científica.</i>	Siempre	125	36%	350
	Algunas veces	146	42%	
	A menudo	59	17%	
	Rara vez	15	4%	
	Nunca	2	1%	
	Internet	317	91%	350

Análisis Estadístico Univinario

¿Qué tipos de fuentes utiliza?	Libros de textos	195	56%	
	Artículos de revistas	42	12%	
	Biografías	80	23%	
	Otros	7	2%	
¿Qué fuente de información, considera usted confiable para realizar sus trabajos de investigación académica?	Rincón del vago	42	13%	350
	Monografías.com	169	48%	
	SlideShare.net	142	41%	
	BuenasTareas.com	80	23%	
	Otros	101	29%	
¿Cuáles de las Bases de Datos Científicas conoce o utiliza para hacer sus trabajos de Investigación Académica?	Redalyc	20	6%	350
	SciELO	26	7%	
	Dialnet	22	6%	
	Google Académico	288	82%	
	Ninguna	43	12%	
Conoce Usted, si la Institución de Educación Superior en la que estudia, posee Base de Datos Científicas	Si	227	65%	350
	No	123	35%	
¿Qué Base de Datos científica utiliza con mayor frecuencia?	Taylor & Francis	81	23%	350
	Springer	44	13%	
	Gale	53	15%	
	ProQuest	62	18%	
	Otras	46	13%	
¿Por qué motivos usted realiza un trabajo de investigación académica?	Cumplir la tarea	192	55%	350
	Conocimiento propio	151	43%	
	Hacer un buen trabajo	158	45%	
	Obtener una calificación	103	29%	
	Otros	7	2%	
	[0 – 1]	39	11%	350
	[2 – 3]	180	51%	

¿Cuántas Horas semanales dedica usted, para hacer su trabajo de investigación?	[4 – 5]	84	24%	
	[6 – 7]	28	8%	
	Mayor a 8	19	5%	
¿Usted se toma el tiempo necesario para leerla y analizarla detenidamente?	Si	321	92%	350
	No	29	8%	
¿Cree usted que adquirió los conocimientos adecuados acerca del tema que investigó?	Total Acuerdo	145	41%	350
	Parcial Acuerdo	170	49%	
	Indiferente	21	6%	
	Parcial Desacuerdo	11	3%	
	Total Desacuerdo	2	1%	
Considera Usted que la falta de tiempo debido a asuntos personales o de trabajo, es un factor que le impide realizar una mejor investigación	Total Acuerdo	165	47%	350
	Parcial Acuerdo	141	40%	
	Indiferente	33	9%	
	Parcial Desacuerdo	6	2%	
	Total Desacuerdo	5	1%	
¿Considera Usted que el docente de la asignatura que le envía una investigación, le hace un seguimiento de su trabajo?	Total Acuerdo	117	33%	350
	Parcial Acuerdo	133	38%	
	Indiferente	62	18%	
	Parcial Desacuerdo	22	6%	
	Total Desacuerdo	14	4%	
¿Considera Usted que sus trabajos de investigación no aportan nada a su formación profesional?	Total Acuerdo	45	13%	350
	Parcial Acuerdo	78	22%	
	Indiferente	54	15%	
	Parcial Desacuerdo	65	19%	
	Total Desacuerdo	108	31%	
	Total Acuerdo	59	17%	350
	Parcial Acuerdo	75	21%	

<i>¿Considera Usted que todo lo que está en internet se puede copiar?</i>	Indiferente	58	17%	
	Parcial Desacuerdo	91	26%	
	Total Desacuerdo	67	19%	
<i>¿Considera Usted que es importante realizar citas bibliográficas, en un trabajo de investigación?</i>	Total Acuerdo	165	47%	350
	Parcial Acuerdo	111	32%	
	Indiferente	43	12%	
	Parcial Desacuerdo	16	5%	
	Total Desacuerdo	14	4%	
<i>¿Cuánto tiempo le dedica usted a la lectura semanalmente?</i>	[0 – 2]	186	53%	350
	[3 – 5]	125	36%	
	[6 – 8]	19	5%	
	Más de	3	1%	
	No leo	17	5%	
<i>Cuanto tiempo le dedica a la investigación de los trabajos enviados por su profesor.</i>	[0 – 2]	3	1%	350
	[3 – 5]	6	2%	
	[6 – 8]	34	10%	
	Más de	160	46%	
	No leo	147	42%	
<i>¿Sabía usted que copiar un trabajo de otra persona, se considera plagio académico?</i>	Total Acuerdo	266	76%	350
	Parcial Acuerdo	46	13%	
	Indiferente	18	5%	
	Parcial Desacuerdo	8	2%	
	Total Desacuerdo	12	3%	
<i>¿Usted cree que la causa de copiar un trabajo académico por parte de los alumnos es debido a la facilidad que ofrece el internet para encontrar información?</i>	Total Acuerdo	167	48%	350
	Parcial Acuerdo	123	35%	
	Indiferente	40	11%	
	Parcial Desacuerdo	10	3%	
	Total Desacuerdo	10	3%	
	Total Acuerdo	160	46%	350

<i>¿Usted cree que la causa de copiar un trabajo académico por parte de los alumnos es debido a que todo lo que se encuentra en internet se puede utilizar, por ser un acceso público?</i>	Parcial Acuerdo	113	32%	
	Indiferente	51	15%	
	Parcial Desacuerdo	15	4%	
	Total Desacuerdo	11	3%	
<i>¿Usted cree que la causa de copiar un trabajo académico por parte de los alumnos, es debido a que NO sabe cómo realizar trabajos de investigación académica?</i>	Total Acuerdo	134	38%	350
	Parcial Acuerdo	123	35%	
	Indiferente	45	13%	
	Parcial Desacuerdo	25	7%	
	Total Desacuerdo	23	7%	

8.3.3 DESCRIPCIÓN Y TIPO DE CADA UNA DE LAS VARIABLES

Tabla 7: Descripción de variables

Variable	Descripción	Tipo de variable
Género	Variable que se define el género del encuestado, el mismo que puede ser masculino y femenino	Cualitativa
Edad	La variable edad determina los años del encuestado expresado en números.	Cuantitativa
IES	Variable que nos indica a que institución pertenece el encuestado	Cualitativa
Facultad	Variable que nos indica a que facultad dentro de la institución a encuestar pertenece.	Cualitativa
Carrera	Variable que nos determina qué carrera cursa el encuestado	Cualitativa
Semestre/Año	Variable que permite conocer el nivel de estudio del encuestado.	Cuantitativa
Asignaturas tomadas	Variable que nos indica la asignatura tomadas por el encuestado en ese semestre /Año	Cuantitativa
Trabajos de investigación	Variable que nos releva si en la institución que pertenece les mandan trabajos investigativos.	Cualitativa
Más trabajos investigación	Variable que nos releva las materias que les mandan trabajos de investigación.	Cualitativa
Docente PHD	Variable que nos indica si la institución posee docentes PHD.	Cualitativa
Lugar de investigación	Variable que nos indica donde el encuestado realiza sus trabajos de investigación	Cualitativa
Frecuencia de investigación	Cuántas veces realiza los trabajos el encuestado.	Cualitativa
Fuentes a utilizar	Fuentes que el encuestado conoce para la investigación	Cualitativa

8.4 ANÁLISIS ESTADÍSTICO UNIVARIADO

A continuación, hemos tomado en consideración algunas variables que pensamos son las más significativas para nuestro estudio, así como una interpretación didáctica, con la finalidad de que el lector quienes son los estudiantes de la Universidad de Guayaquil, puedan tener un respaldo y un soporte mediante un lenguaje apropiado estadístico, para que puedan realizar sus trabajos.

Recordemos que la Estadística es una ciencia que involucra a casi todas las demás ciencias por lo que es importante que cuando se presenta un trabajo elaborado por los estudiantes, se lo realice en formato de informe estadístico, así como la interpretación de los resultados en relación a los contenidos estadísticos, ya que es un aspecto muy importante de evaluar, es significativo que el estudiante comprenda los cálculos, gráficos y procedimientos realizados y en esencial la interpretación de los resultados que ofrece el software (Excel), que se ha utilizado en el contexto de cada uno de los casos de estudios propuestos en este libro.

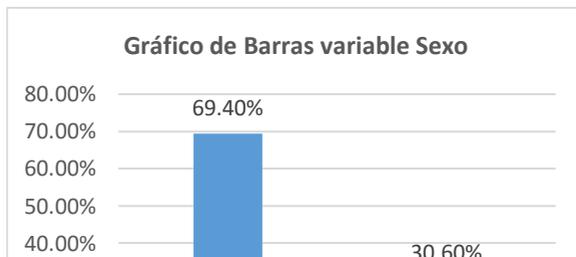
Para este análisis se tomaron en cuenta algunas de las variables correspondientes al trabajo de campo que se realizó, mediante la aplicación de una encuesta, para ello hemos considerado dos tipos de variables, la primera variable a considerar es de tipo cualitativa tanto nominal como ordinal, así mismo tenemos variables de tipo cuantitativa por intervalos también llamadas variables continuas y las variables discretas.

8.4.1 Variable, Género del entrevistado

Tabla 8: Estadística de Género

Género	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Femenino	243	69,4	69,4	69,4
Masculino	107	30,6	30,6	100,0
Total	350	100,0	100,0	

Gráfico 1: Género



Análisis de la variable Género

Este tipo de variable a analizar es de tipo cualitativa nominal, la cual tiene dos categorías, por lo que podemos decir que es una variable dicotómica. Al observar la tabla de frecuencias de esta variable podemos analizar que, de un total 350 estudiantes encuestados que pertenecen a la Universidad de Guayaquil y la Universidad Laica Vicente Rocafuerte, hemos llegado a la conclusión que, 243 estudiantes son de género femenino y que 107 estudiantes encuestados son de género masculino, estos datos son confirmados con el Gráfico 1, por lo que podemos observar que el porcentaje del género femenino corresponde al 69,40% del total de la muestra y que el porcentaje del género masculino es del 30,60%, lo que concluimos, referente a la prevalencia del plagio académico en los trabajos de investigación, que la mayoría de las personas encuestadas correspondieron al sexo femenino.

8.4.2 Variable. Edad del entrevistado (en años).

Tabla 9: Resumen Estadístico.

Edad		
N	Válido	350
	Perdidos	0
Media		22,37
Mediana		22
Moda		21
Desviación estándar		3,289
Varianza		10,819
Asimetría		1,823
Curtósis		5,586
Rango		24
Mínimo		18

Tabla 10: Estadística de Edad

Clase (edad)	Marca de clase	Frecuencia. Absoluta	Frecuencia. Relativa (%)	Frecuencia. Acumulada	F. Relativa acumulada
[17 - 20)	18,5	46	0,131428571	46	0,131428571
[20 - 23)	21,5	179	0,511428571	225	0,642857142
[23 - 26)	24,5	81	0,231428571	306	0,874285714
[26 - 29)	27,5	20	0,057142857	326	0,931428571
[29 - 32)	30,5	18	0,051428571	344	0,982857142
[32 - 35)	33,5	3	0,008571429	347	0,991428571
[35 - 38)	36,5	1	0,002857143	348	0,994285714
[38 - 41)	39,5	1	0,002857143	349	0,997142857
[41 - 44)	42,5	1	0,002857143	350	1
		350	1		

Análisis descriptivo de la variable Edad:

Con la finalidad de poder realizar este análisis, lo hemos hecho, mediante el uso de la tabla de frecuencias y estadística descriptiva

1. Interpretación de datos mediante Tabla de Frecuencias:

La tabla 10 corresponde a una tabla de frecuencias. En dicha tabla podemos observar que el número n de datos es de 350 y que la variable toma valores continuos, definida en 9 intervalos y cuyo ancho de cada intervalo es de 3 elementos incluidos en la muestra y 1 no incluido (por ejemplo, en el primer intervalo tenemos las edades 17,18,19, incluidos y 20 no incluido). Es importante tener claro la diferencia que existe entre los valores correspondientes a la clase, ya que son los valores de la variable edad, de aquellas personas que se entrevistaron, pero en un cierto intervalo de edades (por ejemplo: [17 - 20), [20 - 23)... [41 - 44)); con los valores de la frecuencia absoluta (por ejemplo: 46, 179, ... , 1), ya que si analizamos el primer intervalo, podemos decir que de un total de 350 personas consideradas en este estudio, las edades correspondientes entre 17 y 20 años, 46 corresponden a ese intervalo de edades, esto supone un 13% de 350, por lo que 0.1314 o 13% es la frecuencia relativa del valor correspondiente al intervalos de las edades entre 17 y 20 años, este último valor sin incluir.

Otros elementos que debemos considerar dentro de la tabla de frecuencia son: la marca de clase y la frecuencia relativa

acumulada. Para una mejor interpretación lo haremos a través del gráfico del polígono de frecuencia y de la Ojiva, respectivamente.

El gráfico del polígono de frecuencia y la ojiva son gráficos bidimensionales, En este caso el polígono de frecuencias refleja en el eje de las X las edades de las personas encuestadas (marca de clase) y en el eje de las Y, refleja porcentualmente la cantidad de personas encuestadas (frecuencia relativa), por otro lado, la ojiva representa en el eje de las X la característica a ser estudiada (edad), y en el eje de las Y, el porcentaje acumulado (frecuencia relativa acumulada). Para una mejor interpretación, haremos unas preguntas que nos ayudarán a entender como analizar estos gráficos:

¿Cuál es el promedio aproximado de personas encuestadas cuya edad corresponden a 21,5 años? El promedio aproximado es de 0,5114 es decir el 51,1%

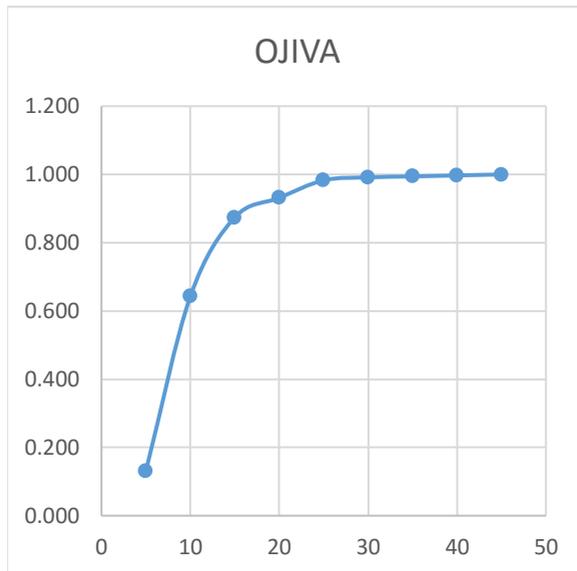
¿Cuál es el promedio acumulado de personas encuestadas, correspondientes a edades entre 18,5 y 30, 5 años de edad? El promedio aproximado es el 98,22%, es decir $(0,1314 + 0,5114 + 0,2314 + 0,057 + 0,051)$ o es lo mismo que $(13,14\% + 51,14\% + 23,14\% + 5,7\% + 5,1\%)$

¿Cuál es el promedio acumulado de personas encuestadas, correspondientes a edades entre 33,5 y 42, 5 años de edad? El promedio aproximado es el 1,69%, es decir $(0,0085 + 0,0028 + 0,0028 + 0,0028)$ o es lo mismo que $(0,85\% + 0,28\% + 0,28\% + 0,28\%)$, esto nos da a entender que la cantidad de personas encuestadas comprendidas en estas edades fueron pocas.

Gráfico 3: Polígono de Edad.



Gráfico 4: Ojiva de Edad.



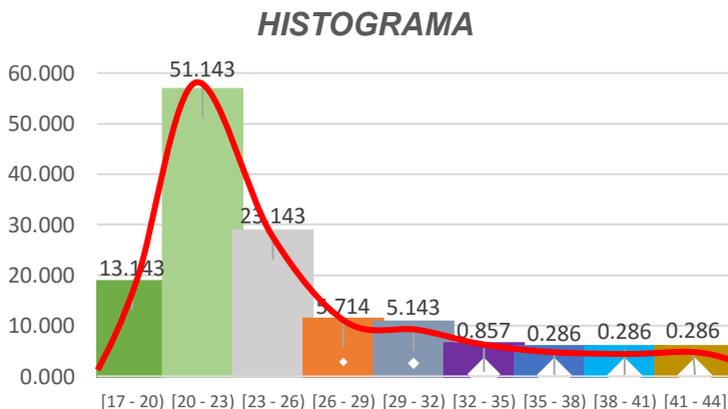
2. Interpretación de datos mediante Estadística descriptiva:

Para el análisis descriptivo de esta variable cuantitativa, debemos tener claro que este tipo de variable nos proporciona 4 grupos de estimadores o estadísticos, el primer grupo corresponde a los estimadores de centralización; Entre los parámetros de centralización, también llamados de tendencia central tenemos tres como los más destacables, que son la media la mediana y la moda. Al analizar el contenido de la tabla 9, podemos indicar que de 350 personas que se entrevistaron para determinar la prevalencia del plagio académico dentro de los trabajos de investigación, han tenido una edad promedio de 22,37 años. La mediana nos indica que la mitad de las personas encuestadas fueron superior de 22 años y la otra mitad fueron inferiores a 22 años, y el valor de la moda nos indica que en su mayoría las personas encuestadas fueron jóvenes de 22 años, es decir el valor que más se repite de la edad dentro de la muestra.

El segundo grupo de estimadores corresponde a los de dispersión, entre ellos tenemos a la varianza, y la desviación estándar, con respecto al valor de la desviación estándar, nos indica que en media los valores de la variable edad se desvían poco más de $\pm 3,289$ años de su promedio. Es decir, la distancia media a la que se sitúan los valores de las edades de los encuestado respecto de su promedio. Con respecto a la varianza es el cuadrado de la desviación estándar y tiene la misma interpretación, solo que nos da una cuantía de la dispersión general, cuyo valor es de 10,819.

El tercer grupo de estimadores corresponde a los de forma, entre los que tenemos al coeficiente de simetría y a la Curtósis, para este caso los valores correspondientes a ambos estimadores son positivos (tabla 9). Con respecto a la forma de la simetría, podemos deducir estadísticamente que la distribución es asimétrica hacia la derecha (recordemos que la asimetría lo da la cola de la distribución), es decir que los datos se encuentran en su mayoría agrupados hacia la izquierda de la distribución, esto lo podemos observar con más claridad en el gráfico adjunto.

Gráfico 5: Histograma de Edad



Autor: Varios

Fuente: Excel Cuadro Estadístico

8.4.3 Variable: “Indique el nombre de la Institución de Educación Superior (IES) a la que Pertenece”.

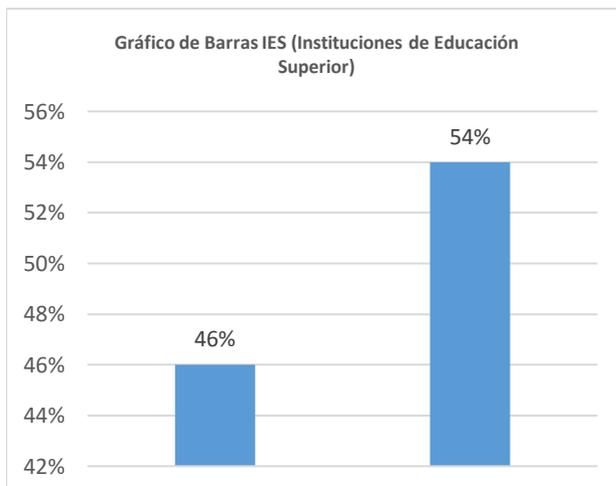
En la tabla 11, tenemos que, de un total de 350 personas encuestadas, que dieron su opinión respecto a la prevalencia del plagio académico en los trabajos de investigación, 161 personas pertenecen a la universidad de Guayaquil y 189 personas pertenecen a la Universidad Laica Vicente Rocafuerte.

Tabla 11: Tabla de frecuencia de IES

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Universidad De Guayaquil	161	46,0	46,0	46,0
Universidad Laica Vicente Rocafuerte	189	54,0	54,0	100,0
Total	350	100,0	100,0	

Para analizar el porcentaje de personas encuestadas que respondieron tanto de la universidad de Guayaquil, como de la Universidad Laica procedemos hacerlo mediante un diagrama de Barras (Gráfico 6).

Gráfico 7: Diagrama de Barras de IES



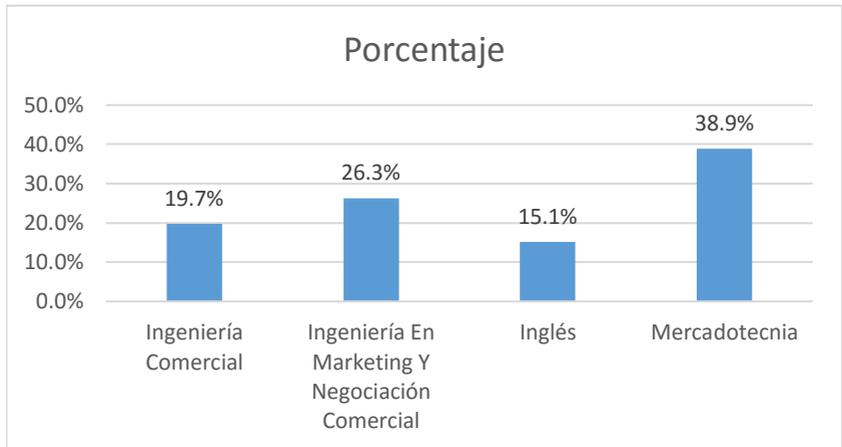
En el Gráfico de Barras IES, observamos que un 46% de estudiantes encuestados de una muestra tomada al azar a 350 estudiantes encuestados son de la Universidad de Guayaquil y un 54% son de la Universidad Laica Vicente Rocafuerte,

8.4.4 Variable: Carrera a la que pertenece el entrevistado

Tabla 43: Tabla de Frecuencia de Carrera

Carrera	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Ingeniería Comercial	69	19,7	19,7	19,7
Ingeniería En Marketing Y Negociación Comercial	92	26,3	26,3	46,0
Inglés	53	15,1	15,1	61,1
Mercadotecnia	136	38,9	38,9	100,0
Total	350	100,0	100,0	

Gráfico 2: Diagrama de barras de Carrera.



Análisis

Este tipo de variable a analizar es de tipo cualitativa nominal, la cual tiene cuatro categorías, por lo que podemos decir que es una variable politómica. Al observar la tabla de frecuencias de esta variable podemos analizar que de una muestra tomada al azar a 350 estudiantes encuestados de la Universidad de Guayaquil y la Universidad Laica Vicente Rocafuerte, hemos llegado a la conclusión que en la Tabla de frecuencias de la variable Carrera, podemos observar que 69 estudiantes encuestados pertenecen a la carrera de Ingeniería Comercial, 92 estudiantes encuestados pertenecen a la carrera de Ingeniería en Marketing y Negociación Comercial, 53 estudiantes encuestados pertenecen a la carrera de Inglés y 163 estudiantes encuestados pertenecen a la carrera de Mercadotecnia. En el Gráfico de Barras de la misma variable, podemos analizar que el porcentaje de estudiantes encuestados

de la Carrera de Ingeniería Comercial es del 19,71%, el porcentaje de estudiantes encuestados de la Carrera Ingeniería en Marketing y Negociación Comercial es del 26,29%, el porcentaje de estudiantes encuestados de la Carrera de Ingles es del 15,14% y que el porcentaje de estudiantes encuestados de la Carrera de Mercadotecnia es del 38,86%.

8.4.5 Variable: Horas semanales dedica para hacer trabajos de investigación

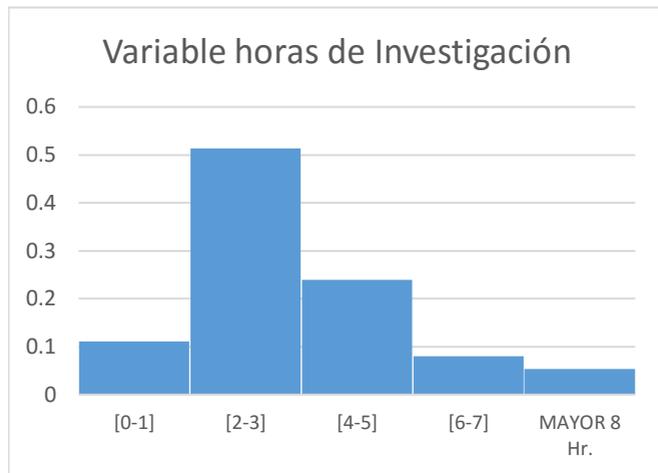
Tabla 20: Resumen Estadístico

Horas investigación		
N	Válido	350
	Perdidos	0
Media		2,45
Mediana		2,00
Moda		2
Desviación estándar		,979
Varianza		,959
Asimetría		,910
Curtosis		,622
Rango		4
Mínimo		1
Máximo		5
Percentiles	25	2,00

Tabla 21: Tabla de frecuencia de Horas_ investigación

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	[0-1]	39	11,1	11,1	11,1
	[2-3]	180	51,4	51,4	62,6
	[4-5]	84	24,0	24,0	86,6
	[6-7]	28	8,0	8,0	94,6
	MAYOR 8 Hr.	19	5,4	5,4	100,0
	Total	350	100,0	100,0	

Gráfico 5: Histograma de Horas de investigación



Análisis

En el resumen estadístico, para la variable Horas semanales dedicada hacer trabajo de investigación, nos podemos dar cuenta que es una variable cuantitativa continua, es decir se encuentra estructurada mediante intervalos; de acuerdo a la frecuencia de cada intervalo podemos interpretar (tabla 21), que de un total de

350 estudiantes elegidos al azar, tanto de la universidad de Guayaquil, como de la Universidad Laica, 39 estudiantes dedican como máximo 1 hora diaria para investigar, 180 estudiantes entre 2 y 3 horas, 84 estudian ente 4 y 5 horas, 28 entre 6 y 7, y apenas 19 estudiantes más de 8 horas, esto nos da a pensar de que existen muy pocas personas que se dedican a investigar, y es más probable que la mayoría de los encuestados cometa algún tipo de plagio en sus trabajos, debido a las pocas horas que le dedican a la investigación, es decir un promedio de 2.45 horas. Con respecto al valor de la mediana nos da 2 horas, si nos ponemos a analizar estos dos estadísticos, es decir la media y la mediana, podemos damos cuenta que la media es mayor, esto nos da a interpretar otro estadístico que es el coeficiente de simetría de Fisher, por lo que podemos asegurar de que los datos correspondientes al gráfico del histograma (26), presentan una asimetría hacia la derecha, esto implica que en su mayoría los datos se acumulan hacia la izquierda de la distribución, lo que nos da a indicar que en su mayoría las personas investigan un máximo de 2 horas para hacer sus trabajo académicos de investigación.

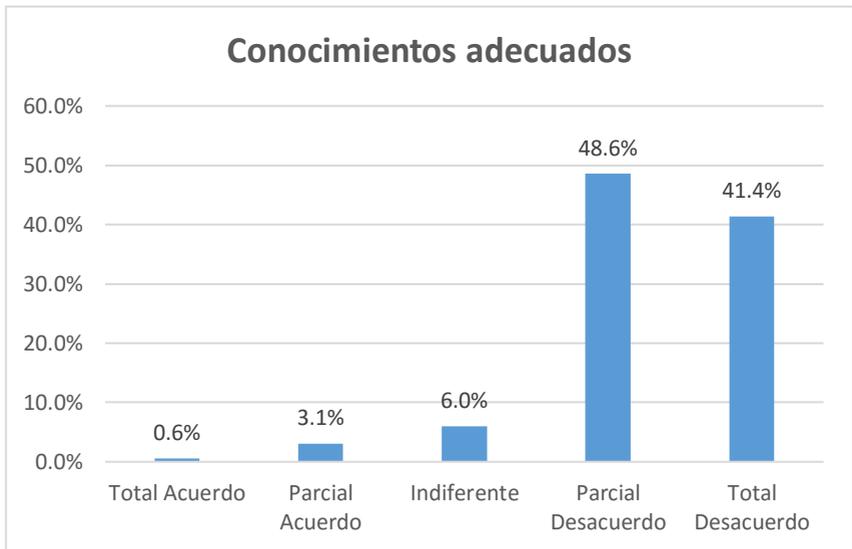
Con respecto a la Curtósis (tabla 20), podemos indicar que el grado de concentración que presentan los valores alrededor de la media es mayor a cero (0.622) por lo que la distribución es Leptocúrtica. Con respecto al valor de la desviación estándar podemos decir que la distancia de cada punto respecto a la media es de 2.45 ± 0.979 horas de lectura.

8.4.6 Variable: Cuando finaliza su trabajo de investigación ¿Cree usted que adquirió los conocimientos adecuados acerca del tema que investiga?

Tabla 6: Tabla de Frecuencia

Conocimiento Adecuado	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Total, Acuerdo	2	0,6	0,6	0,6
Parcial Acuerdo	11	3,1	3,2	3,7
Indiferente	21	6	6	9,7
Parcial Desacuerdo	170	48,6	48,7	58,5
Total,	115	11,4	11,5	100

Gráfico 3: Diagrama de barras



Análisis

Este tipo de variable a analizar es de tipo cualitativa ordinal, es decir tiene un orden de prioridad, si nos damos cuenta esta variable tiene 5 categorías. Al observar la tabla de frecuencias y el gráfico de barras podemos analizar que de una muestra tomada al azar a 350 estudiantes encuestados de la Universidad de

Guayaquil y la Universidad Laica Vicente Rocafuerte, el 90% están en parcial desacuerdo y total desacuerdo que cuando hacen una investigación o una tarea que tenga que ver con la investigación, no adquieren los conocimientos adecuados de la investigación que realizo, si nos ponemos a inferir, podríamos decir que el alumno simplemente copia y pega la investigación y no se toma el tiempo de leerla o repasarla.

9 ANÁLISIS ESTADÍSTICO DESCRIPTIVO APLICADO A LA DESERCIÓN ESTUDIANTIL DE LOS CURSOS DE NIVELACIÓN, CASO DE ESTUDIO: UNIVERSIDAD DE GUAYAQUIL

Una de las dificultades en el proceso de enseñanza aprendizaje, analizada desde el punto de vista pedagógico, es la necesidad de formación de las nuevas generaciones de estudiantes y el compromiso de que cumplan con el papel de asegurar la reproducción de la estructura social de clases en una formación determinada, es por esa razón que al realizar una entrevista con la coordinadora general de los programas de nivelación de la universidad de Guayaquil, nos dio a conocer que durante todo este proceso, que ya hasta hoy es el quinto que se ha realizado, existen algunas variables claves que me ayudan a identificar mediante una encuesta a los estudiantes los factores más relevantes para que exista una deserción.

Durante todo el proceso de análisis, referente a este tema se presentan algunas definiciones estadísticas necesarias para su comprensión, también se describen los tipos de muestreo a utilizar, se presenta el marco muestral, el diseño muestral y la determinación del tamaño de la muestra respectivamente, también se describe el diseño del cuestionario y las diferentes secciones que lo componen, se presenta la descripción y codificación de las variables a utilizarse.

9.1 Población Objetivo.

Siendo la Población Objetivo el conjunto de entes cuyas características se desea investigar, para este caso se determina como población objetivo los estudiantes del curso de nivelación de una universidad de Guayaquil, correspondientes a las Facultades de Ciencias Económicas, Ciencias Administrativas,

Filosofía, letras y ciencias de la educación, Jurisprudencia, Ciencias Matemáticas & Físicas y Psicología, en los períodos comprendidos marzo 2014 - agosto 2014. Se basa fundamentalmente en determinar los posibles factores de la deserción estudiantil, que hace que el estudiante abandone los cursos de nivelación propuestos por el gobierno nacional.

9.2 Marco Muestral.

El marco muestral es la representación o codificación de los entes que constituyen la población objetivo. Para esta investigación el marco muestral corresponde a una base de datos de estudiantes matriculados en el curso de nivelación de diecisiete facultades que tiene la Universidad de Guayaquil, de las cuales se han tomado para nuestro estudio a seis de ellas (Ciencias Económicas, Ciencias Administrativas, Filosofía letras y ciencias de la educación, Jurisprudencia, Ciencias Matemáticas & Físicas y Psicología), en el periodo comprendido entre febrero del 2012 a agosto del 2014, proporcionada por el departamento de coordinación general de los cursos de nivelación.

La Universidad de Guayaquil, cuenta con diecisiete facultades, pero para nuestro análisis solo trabajaremos con seis facultades (Ciencias Económicas, Ciencias Administrativas, Filosofía, letras y ciencias de la educación, Jurisprudencia, Ciencias Matemáticas & Físicas y Psicología), cada una de ellas considerada un estrato con una ponderación de 0.107, 1.162, 0.149, 0.1414, 0.253, 0.185 respectivamente, determinada por la proporción de estudiantes que ingresaron a cada carrera durante el periodo en estudio. El tipo de muestreo a utilizar es el muestreo bietápico. Primera etapa: Muestreo estratificado; siendo los estratos las seis diferentes facultades de una universidad pública. Segunda etapa: Muestreo aleatorio simple.

En el Tabla# 5 se presenta los tamaños poblacionales correspondientes a cada carrera.

Tabla 7: Carreras

CARRERAS	Tamaños Poblacionales
CIENCIAS ECONOMICAS	N1 = 130
FACULTAD DE ADMINISTRACION	N2 = 197
FILOSOFÍA, LETRAS Y CIENCIAS DE LA EDUCACIÓN	N3 = 181
JURISPRUDENCIA	N4 = 171
MATEMATICA – FISICA	N5 = 306
PSICOLOGIA	N6 = 224
Total	1209

9.3 Determinación del tamaño de la muestra.

Para nuestro análisis hemos tomado como el tamaño de la muestra el mismo tamaño de la población objeto de estudio, esto con la finalidad de tener una mayor confiabilidad en la realización de este trabajo, el tamaño correspondiente fue de 1211 alumnos encuestados.

9.4 Instrumento de Recolección de Datos

Mediante una encuesta realizada en la que se tomaron al azar seis facultades de un total de diecisiete facultades que tiene la Universidad de Guayaquil. Es importante recalcar que esta encuesta se la realizo la primera semana del inicio del curso de nivelación correspondiente al primer semestre del periodo de clases del curso de nivelación del 2014 (1s-2014), es decir marzo 2014 - agosto del 2014.

Esta encuesta se la realizo a los estudiantes de aquellas facultades escogidas al azar, con la finalidad de conocer su percepción con respecto a las posibles causas que ellos tendrían, como motivo para tomar la decisión de desertar; ahora para corroborar y cotejar los resultados obtenidos, se trabajó en conjunto con el departamento de coordinación general de los cursos de

nivelación, haciendo un seguimiento de estos estudiante llenando una base de datos en Excel, en donde reflejábamos semana a semana la asistencia de los alumnos por facultades esta base era llenada gracias a la información que los coordinadores de cada una de estas facultades nos proporcionaban, este trabajo se lo hacía con la finalidad de monitorear el momento en que el estudiante ya no asistía a clases y gracias a esta información podíamos detectar en que semana desertaba.

También se pudo encuestar a estudiantes que habían desertado; esto se pudo lograr, gracias a la ayuda de los tutores de los cursos de nivelación, correspondientes a cada una de las facultadas elegidas al azar, quienes poseían una ficha estudiantil donde se encontraban los datos personales de cada uno de los alumnos por lo que se hizo un poco más fácil el contactarlos y así poderles hacer la encuesta, el total de alumnos que se pudo contactar fue de 105 alumnos.

9.5 Diseño del cuestionario

El cuestionario está distribuido en cinco secciones, la primera denominada “Información General del Entrevistado”, que consta de ocho preguntas donde se obtiene información tal como: género, fecha ciudad donde vive, estado civil, etc. La segunda sección tiene como título “Situación Laboral del Estudiante”, en esta sección que consta de dos preguntas, se busca obtener información de su situación actual en el aspecto laboral, así como el tiempo que le dedica a su trabajo, en la siguiente sección denominada “Información Familiar” consta de siete preguntas y se busca conocer con quien vive el alumno, que tipo de vivienda tiene los niveles educativos del padre y de la madre, así como el número de hermanos que tiene. En la cuarta sección “Información Académica” se presentan tres preguntas en las se busca obtener información que nos proporcione una idea del desarrollo académico del entrevistado durante sus estudios básicos. Y finalmente en la última sección se presenta un formulario de

“Causas posibles de Deserción” mediante el cual se obtendrá información sobre los posibles factores que lo pueden llevar o lo llevaron a tomar la decisión de desertar de sus estudios

9.5.1 ENCUESTA

1. INFORMACIÓN GENERAL

1.1 Género: Masculino ___ Femenino ___

1.2 Edad _____

1.3 Estado Civil: Soltero ___ Casado ___ Unión Libre ___ Otro ___ ¿Cuál? _____

1.4 Ciudad donde vive _____

1.5 Tipo de colegio donde cursó sus estudios: Privado ___ Público ___

1.6 Jornada en la que estudió: Matutina ___ Vespertina ___ Nocturna ___

1.7 El tipo de Bachillerato que usted está cursando es:

Bachillerato General Unificado ___ Bachillerato en Ciencias ___ Bachillerato Técnico ___ Bachillerato Internacional ___ otro ___ ¿Cuál? _____

1.8 Como solventa sus necesidades económicas:

Ingresos Propios ___ Depende de sus padres ___ Otro ___ ¿Cuál? _____

2. SITUACIÓN LABORAL DEL ESTUDIANTE

2.1 Usted Trabaja: Si ___ No ___

2.2 tiempo que le dedica al trabajo: Tiempo Completo ___ Medio Tiempo ___ Tiempo Parcial ___ Otro ___ ¿Cuál? _____

3. INFORMACIÓN FAMILIAR

3.1 ¿Con quién vive? _____

3.2 La vivienda donde vive es: Propia ___ Arrendada ___ Cedida _____

3.3 Tipo de Vivienda: Casa ___ Departamento ___ Villa ___ Edificio ___ Cuarto _____

3.4 Estructura de la Vivienda: Cemento ___ Madera ___ Mixta _____

3.5 Número de Hermanos que tiene: _____

3.6 Nivel educativo del Padre: Artesano ___ Bachiller ___ Profesional ___ Otro ___ ¿Cuál? _____

3.7 Nivel educativo de la Madre: Artesano ___ Bachiller ___ Profesional ___ Otro ___ ¿Cuál? _____

4.- INFORMACIÓN ACADÉMICA

4.1 Culminado su bachillerato, cuál fue su opción de UNIVERSIDAD PÚBLICA:
ESPOL: ____ Universidad de Guayaquil: ____ Otro ____ ¿Cuál?

4.2 Cuales fueron sus primeras dos opciones de carrera al momento de llenar el formulario Web de ingreso a las universidades públicas:

a: _____

b: _____

4.3 La carrera que usted tenía como prioridad para su vida profesional fue la misma que puso como opción en el formulario web de ingreso a las universidades públicas.

Sí ____ No ____

4.4 Al momento de Ingresar al curso de nivelación de la carrera asignada, usted conocía el Plan de Estudio: Sí ____ No ____

4.5 Al momento de Ingresar al curso de nivelación de la carrera asignada, usted conocía el Campo de Trabajo donde se desempeñaría al momento de egresar: Sí ____ No ____

4.6 Al momento de Ingresar al curso de nivelación de la carrera asignada, usted tenía conocimientos previos de **HABILIDADES DE RAZONAMIENTO DEL PENSAMIENTO ABSTRACTO**:

Sí ____ No ____

4.7 Al momento de Ingresar al curso de nivelación de la carrera asignada, usted tenía conocimientos previos de **INTRODUCCIÓN A LA COMUNICACIÓN CIENTÍFICA**: Sí ____ No ____

4.8 Al momento de Ingresar al curso de nivelación de la carrera asignada, usted tenía conocimientos previos de **RESOLUCION ESTRATÉGICA DE PROBLEMAS**: Sí ____ No ____

4.9 Al momento de Ingresar al curso de nivelación de la carrera asignada, usted tenía conocimientos previos de **MATEMÁTICAS**: Sí ____ No ____

5. EN CASO DE QUERERSE RETIRAR DEL CURSO DE NIVELACIÓN O YA SE HA RETIRADO, INDIQUE LAS CAUSAS POR LAS CUALES USTED LO HARÍA O LO HIZO

NOTA: En las Opciones indique el grado de prioridad:

En el caso de Opciones que tienen 3 respuestas marque: 3 como máxima prioridad y *1* como la mínima prioridad

En el caso de Opciones que tienen 6 respuestas marque: 6 como máxima prioridad y *1* como la mínima prioridad

5.1 ECONÓMICOS – LABORALES

- a. Dificultad para cumplir con los gastos adicionales durante todo el curso de nivelación (transporte, copias del material de estudios, etc.).
 - b. Dificultad para cumplir con el horario de estudios por motivos laborales
 - c. Otra Esplique
-

5.2 FAMILIARES

- a. Presión por parte de los familiares para abandonar los estudios
- b. Cambio de Ciudad
- c. Cambio de Estado Civil
- d. Paternidad, Maternidad
- e. Enfermedad
- f. Influencia de personas diferentes al núcleo familiar (Amigos, Compañeros, etc.)

5.3 PSICOLÓGICAS

- a. La carrera no cumplió con sus expectativas
 - b. Su motivación para continuar con sus estudios cambio en el transcurso del curso de nivelación
 - c. Adaptación social con sus compañeros
 - d. Dificultad para adaptarse al cambio del colegio a la universidad
 - e. Inestabilidad familiar y problemas afectivos
 - f. Otra Esplique
-

5.4 INSTITUCIONALES

- a. El sistema académico universitario (horarios, Criterios de evaluación, etc.)
- b. Conflictos vividos con docentes y/o estudiantes

- c. Le gustaba la carrera, pero no la universidad (Planta física, lugar) _____

9.6 DESCRIPCIÓN Y CODIFICACIÓN DE VARIABLES

Para facilitar la interpretación y manejo de las variables investigadas, es de gran utilidad establecer una codificación. A continuación, se presenta la descripción y codificación determinada para cada una de las variables de acuerdo al diseño del cuestionario.

9.6.1 CARACTERÍSTICAS GENERALES DEL ENTREVISTADO

Variable 1: Facultad

La Primera variable nos indica en que facultad se encuentra estudiando el entrevistado

Tabla 1: Facultad

CODIFICACION DE LA VARIABLE: FACULTAD	
CIENCIAS ECONOMICAS	1
FACULTAD DE ADMINISTRACION	2
FILOSOFIA, LETRAS Y CIENCIAS DE LA EDUCACION	3
JURISPRUDENCIA	4
MATEMATICA - FISICA	5
PSICOLOGIA	6

Variable 2: Carrera

La segunda Variable nos indicara la carrera correspondiente en que el encuestado pertenece.

Tabla 2: Carrera

CODIFICACION DE LA VARIABLE: CARRERA

COMERCIO EXTERIOR	1
CONTADOR PUBLICO AUTORIZADO	2
DERECHO	3
ECONOMIA	4
GESTION EMPRESARIAL	5
INGENIERIA CIVIL	6
INGENIERIA COMERCIAL	7
INGENIERIA EN MARKETING	8
INGENIERIA EN SISTEMAS COMPUTACIONALES	9
INGENIERIA EN TRIBUTACION Y FINANZAS	10
INGENIERIA EN SISTEMAS ADMINISTRATIVOS	11
MERCADOTECNIA Y PUBLICIDAD	12
PSICOLOGIA	13
SISTEMAS MULTIMEDIA	14

Variable 3: Genero

La variable Género nos indica si el entrevistado, es hombre o mujer.

Tabla 3: Género

CODIFICACION DE LA VARIABLE: GENERO	
MASCULINO	1
FEMENINO	2

Variable 4: Edad

Esta variable nos dará a conocer la edad que nuestro entrevistado tiene.

Variable 5: estado civil

Esta variable nos dará a conocer el estado civil actual del entrevistado, con la finalidad de realizar proporciones de resultados de acuerdo a su estado civil.

Tabla 4: Estado Civil

CODIFICACIÓN DE LA VARIABLE: ESTADO CIVIL	
CASADO	1
SOLTERO	2
UNION LIBRE	3

Variable 6: Ciudad

Esta variable nos indicara el Lugar de procedencia del entrevistado

Tabla 5: Ciudad

CODIFICACION DE LA VARIABLE: CIUDAD			
BABAHOYO	1	NOBOL	18
BALAO	2	PAJAN	19
BUWAY	3	PEDRO CARBO	20
DAULE	4	PLAYAS	21

DURAN	5	PORTOVIEJO	22
EL EMPLAME	6	QUININDE	23
EL TRIUNFO	7	SALITRE	24
ESPERALDA	8	SAMBORONDON	25
GUAYAQUIL	9	SANTA ELENA	26
JIPIJAPA	10	SANTA LUCIA	27
JUJAN	11	TAMARINDO	28
LAS CAÑAS	12	TRIUNFO	29
LIBERTAD	13	TRONCAL	30
MACHALA	14	VENTANAS	31
MILAGRO	15	VILLAMIL PLAYAS	32
NARANJAL	16	YAGUACHI	33
NARANJITO	17		

Variable 7: TIPO COLEGIO

Con esta séptima variable obtendremos el tipo de colegio el cual pertenece el entrevistado.

Tabla 6: Tipo Colegio

CODIFICACION DE LA VARIABLE: TIPO COLEGIO	
PRIVADO	1
PUBLICO	2

Variable 8: JORNADA ESTUDIO

Variable característica para definir y diferenciar la jornada en la que el entrevistado se encuentra estudiando.

Tabla 7: Jornada de Estudio

CODIFICACION DE LA VARIABLE: JORNADA ESTUDIO	
MATUTINA	1
VESPERTINA	2
NOCTURNA	3

Variable 9: TIPO BACHILLERATO

Con esta variable conoceremos en que especialidad se graduó el entrevistado.

Tabla 8: Tipo Bachillerato

CODIFICACION DE LA VARIABLE: TIPO BACHILLERATO	
GENERAL UNIFICADO	1
CIENCIAS	2
TECNICO	3
INTERNACIONAL	4
OTRO	5

Variable 10: SOLVENCIA ECONOMICA.

Esta variable es para saber qué apoyo económico el estudiante recibe para solventar sus estudios

Tabla 9: Solvencia Económica

CODIFICACION DE LA VARIABLE: SOLVENCIA ECONOMICA	
PROPIOS	1
PADRES	2
OTROS	3

9.6.2 SITUACIÓN LABORAL DEL ESTUDIANTE

Variable 11: TRABAJA

La variable 11 como característica general queremos saber el estado actual del entrevistado con respecto a su solvencia económica

Tabla 10: Trabaja

CODIFICACION DE LA VARIABLE: TRABAJA	
SI	1
NO	0

Variable 12: TIEMPO DEDICACION AL TRABAJO

Con esta variable obtendremos una visión del comportamiento y empeño del entrevistado en su ámbito laboral.

Tabla 11: Tiempo dedicado al trabajo

CODIFICACION DE LA VARIABLE: TIEMPO DEDICACION AL TRABAJO	
COMPLETO	1
MEDIO TIEMPO	2
PARCIAL	3
OTRO	4

9.6.3 INFORMACION FAMILIAR

Variable 13: VIVE CON

Mediante esta variable se puede saber con quién vive el entrevistado durante el periodo como estudiante. Es decir, con esta Variable se busca conocer si el entrevistado vivía dentro de un núcleo familiar formado o no.

Tabla 12: Vive Con

CODIFICACION DE LA VARIABLE: VIVE CON	
ABUELOS	1
AMIGOS	2
CONYUGUE	3

HERMANOS	4
HIJOS	5
MADRE	6
PADRASTRO	7
PADRES	8
PADRINOS	9
PAPA	10
PRIMOS	11
SOLO	12
SUEGROS	13
TIOS	14

Variable 14: LA VIVIENDA ES

Mediante esta variable se puede saber el ámbito social y económico el cual presenta el entrevistado.

Tabla 13: La Vivienda es

CODIFICACION DE LA VARIABLE: LA VIVIENDA ES	
PROPIA	1
ARRENDADA	2
CEDIDA	3

Variable 15: TIPO VIVIENDA

Mediante esta variable se puede saber el tipo de vivienda que el entrevistado habita.

Tabla 14: Tipo de Vivienda

CODIFICACION DE LA VARIABLE: TIPO VIVIENDA	
CASA	1

DEPARTAMENTO	2
VILLA	3
EDIFICIO	4
CUARTO	5

Variable 16: ESTRUCTURA VIVIENDA

Mediante esta variable se puede saber la estructura y formación de la vivienda del entrevistado.

Tabla 15: Estructura de la Vivienda

CODIFICACION DE LA VARIABLE: ESTRUCTURA VIVIENDA	
CEMENTO	1
MADERA	2
MIXTA	3

Variable 17: NUMERO HERMANOS

Mediante esta variable se puede saber el número de hermanos del entrevistado

Variable 18: NIVEL EDUCATIVO PADRE

Con esta variable se quiere dar a conocer cuál es el nivel educativo del padre.

ç

Tabla 16: Nivel Educativo del Padre

CODIFICACION DE LA VARIABLE: NIVEL EDUCATIVO PADRE	
ARTESANO	1

BACHILLER	2
PROFESIONAL	3
OTRO	4

Variable 19: NIVEL EDUCATIVO MADRE

Con esta variable se quiere dar a conocer cuál es el nivel educativo de la madre.

Tabla 17: Nivel Educativo de la Madre

CODIFICACION DE LA VARIABLE: NIVEL EDUCATIVO MADRE	
ARTESANO	1
BACHILLER	2
PROFESIONAL	3
OTRO	4

9.6.4 INFORMACION ACADEMICA

Variable 20: OPCION UNIVERSIDAD

Con esta Variable se quiere llegar a conocer las opciones de universidades públicas que fueron seleccionadas previos al ingreso a la universidad.

Tabla 18: Opción Universidad

CODIFICACION DE LA VARIABLE: OPCION UNIVERSIDAD	
ESPOL	1
GUAYAQUIL	2
OTRO	3

Variable 21: PRIMERA OPCION CARRERA

Con esta Variable se llega a conocer la primera opción de carrera al momento de llenar el formulario web de ingreso a la universidad pública.

Variable 22: CARRERA PRIORIDAD

Con esta Variable se llega a conocer la prioridad que tuvo al escoger una carrera es la misma que la que actualmente está cursando.

Tabla 19: Carrera Prioridad

CODIFICACION DE LA VARIABLE: CARRERA PRIORIDAD	
SI	1
No	0

Variable 23: CONOCIMIENTO PLAN DE ESTUDIO

Con esta Variable se llega a conocer si el estudiante conocía el plan de estudio de la carrera que seleccionó.

Tabla 20: Conocimiento Plan de Estudios

CODIFICACION DE LA VARIABLE: CONOCIMIENTO PLAN DE ESTUDIO	
SI	1
No	0

Variable 24: CONOCIMIENTO CAMPO DE TRABAJO

Esta variable nos ayuda a saber si el estudiante tenía conocimiento del campo del trabajo con respecto a la carrera que eligió.

Tabla 21: Conocimiento Campo de Trabajo

CODIFICACION DE LA VARIABLE: CONOCIMIENTO CAMPO DE TRABAJO	
SI	1
No	0

Variable 25: CONOCIMIENTO RAZONAMIENTO AL PENSAMIENTO ABSTRACTO

Esta variable nos ayuda a saber si el estudiante tenía previos conocimientos de la asignatura de razonamiento abstracto previo al ingreso a la universidad.

Tabla 22: Conocimiento Razonamiento Abstracto

CODIFICACION DE LA VARIABLE: RAZONAMIENTO PENSAMIENTO ABSTRACTO	
SI	1
No	0

Variable 26: CONOCIMIENTO COMUNICACIÓN CIENTÍFICA

Esta variable nos ayuda a saber si el estudiante tenía previos conocimientos de la asignatura de comunicación científica previo al ingreso a la universidad.

Tabla 23: Conocimiento Comunicación Científica

CODIFICACION DE LA VARIABLE: CONOCIMIENTO COMUNICACIÓN CIENTÍFICA	
SI	1
No	0

Variable 27: CONOCIMIENTO RESOLUCIÓN DE PROBLEMAS

Esta variable nos ayuda a saber si el estudiante tenía previos conocimientos de la asignatura de resolución de problemas previo al ingreso a la universidad.

CODIFICACION DE LA VARIABLE: CONOCIMIENTO RESOLUCIÓN DE PROBLEMAS	
SI	1

No	0
----	---

Tabla 24: Conocimiento Resolución de Problemas

Variable 28: CONOCIMIENTO MATEMÁTICAS

Esta variable nos ayuda a saber si el estudiante tenía previos conocimientos de la asignatura de matemáticas previo al ingreso a la universidad.

Tabla 25: Conocimiento de Matemáticas

CODIFICACION DE LA VARIABLE: CONOCIMIENTO MATEMÁTICAS	
SI	1
No	0

9.6.5 INFORMACION RESPECTO A LA CAUSA DE DESERCIÓN

Variable 29: CAUSAS ECONOMICAS LABORALES

Esta variable nos ayuda a conocer si la causa de deserción que tuvo el alumno fue debido a problemas económicos laborales.

Tabla 26: Causas Económicas Laborales

CODIFICACION DE LA VARIABLE: CUASAS ECONÓMICAS LABORALES	
Cumplir gastos	1
Cumplir Horarios	2
Otro	3

Variable 30: CAUSAS FAMILIARES

Esta variable nos ayuda a conocer si la causa de deserción que tuvo el alumno fue debido a problemas familiares.

Tabla 27: Causas Familiares

CODIFICACION DE LA VARIABLE: CUASAS FAMILIARES	
Presión Familiares	1
Cambio Ciudad	2
Cambio Estado Civil	3
Paternidad - Maternidad	4
Enfermedad	5
Influencia otros familiares	6

Variable 31: CAUSAS PSICOLÓGICAS

Esta variable nos ayuda a conocer si la causa de deserción que tuvo el alumno fue debido a causas Psicológicas.

Tabla 28: Causas Psicológicas

CODIFICACION DE LA VARIABLE: CUASAS PSICOLOGICAS	
Carrera sin expectativas	1
Cambio motivación	2
Adaptación Social	3
Adaptación Institucional	4
Inestabilidad Familiar	5
Otras causas	6

Variable 32: CAUSAS INSTITUCIONALES

Esta variable nos ayuda a conocer si la causa de deserción que tuvo el alumno fue debido a causas Institucionales.

Tabla 29: Causas Institucionales

CODIFICACION DE LA VARIABLE: CUASAS INSTITUCIONALES	
Sistema Académico	1
Conflictos profesores	2
No le Gusta universidad	3

9.7 Metodologías estadísticas para identificar los factores asociados a la deserción estudiantil.

Para analizar los posibles factores predictivos que ayudarían a explicar las causas de la deserción estudiantil, se ha tomado en consideración realizarlo mediante procedimientos y técnicas estadísticas tales como, Estadística Descriptiva Univariada.

9.8 INTERPRETACIÓN Y ANÁLISIS DE LOS RESULTADOS

9.8.1 Análisis Univariado de los Datos

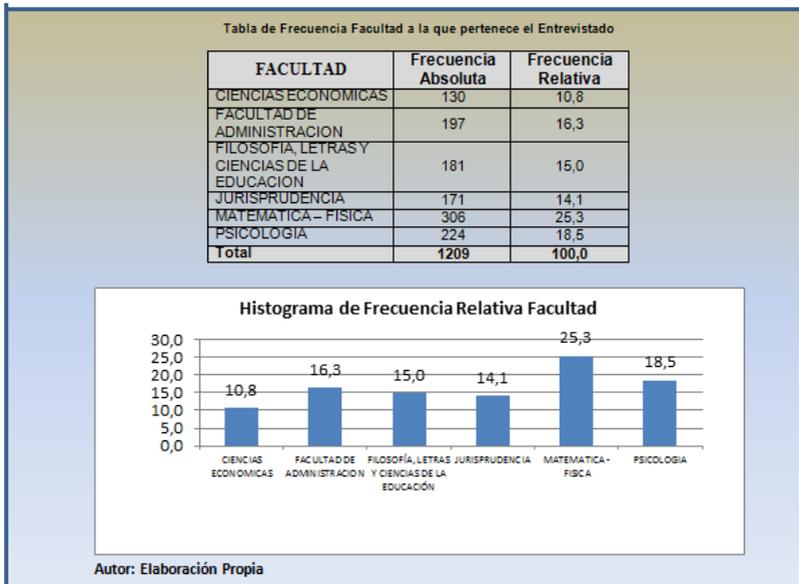
Al realizar un análisis Univariado, este consiste en analizar cada una de las variables que forman parte del estudio en forma separada, es decir el análisis está basado en una sola variable. Para el análisis estadístico de “La deserción estudiantil”, hemos tomado algunas de las variables que consideramos más significativas dentro de nuestro estudio. Es importante recalcar que los datos sobre los que se realiza el análisis, fueron recolectados a través de un cuestionario de preguntas, para ello presentamos los respectivos parámetros de cada variable, además de los gráficos que describen el comportamiento de cada una de las variables que forman parte del estudio.

Para el análisis de las variables cualitativas se procede al cálculo e interpretación de las propiedades correspondientes a cada una de las características que conforman cada una de estas variables. Para el caso de las variables cuantitativas se procederá al análisis de los estimadores tanto de centralización (media, moda, mediana), dispersión (varianza, desviación estándar, rango y coeficiente de variación), posición (cuartiles, diagrama de caja) y de forma (Sesgo y Curtosis).

9.8.2 SECCIÓN: INFORMACIÓN GENERAL DEL ENTREVISTADO

9.8.2.1 Variable: Facultad donde estudia el entrevistado

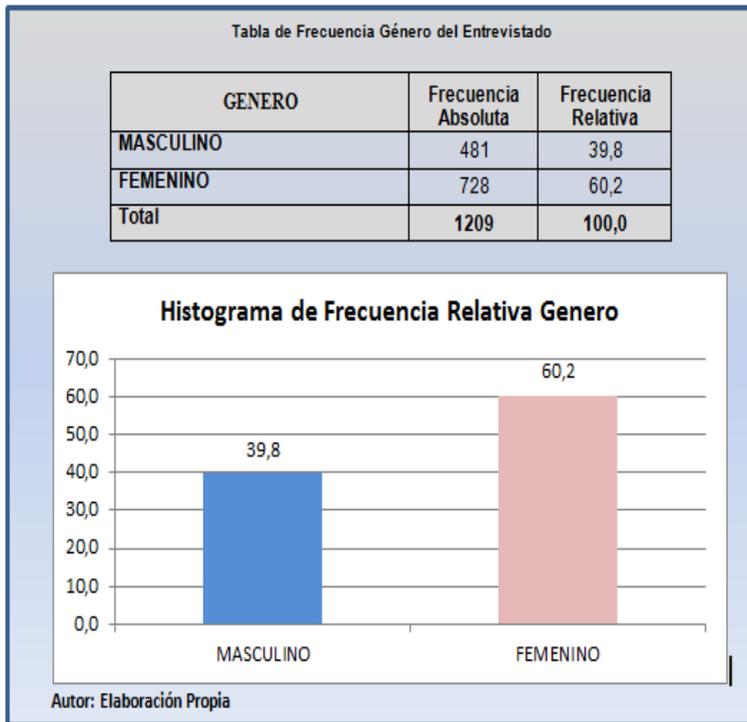
Cuadro 1: Facultad donde estudia el Entrevistado



Esta variable de tipo cualitativa nominal politómica, nos muestra las diferentes facultades en donde se han hecho las encuestas, es así que, de un total de 1209 alumnos encuestados, podemos observar en el histograma de frecuencias (Cuadro 1), que el 25.3% pertenecen a la facultad de Ciencias matemáticas y Físicas, seguido de la facultad de Psicología con un 18.5%, Administración con un 16.3%, Filosofía con un 15%, Jurisprudencia con un 14% y Ciencias Económicas con un 10.8%.

9.8.2.2 Variable: Género del Entrevistado

Cuadro 2: Género del Entrevistado

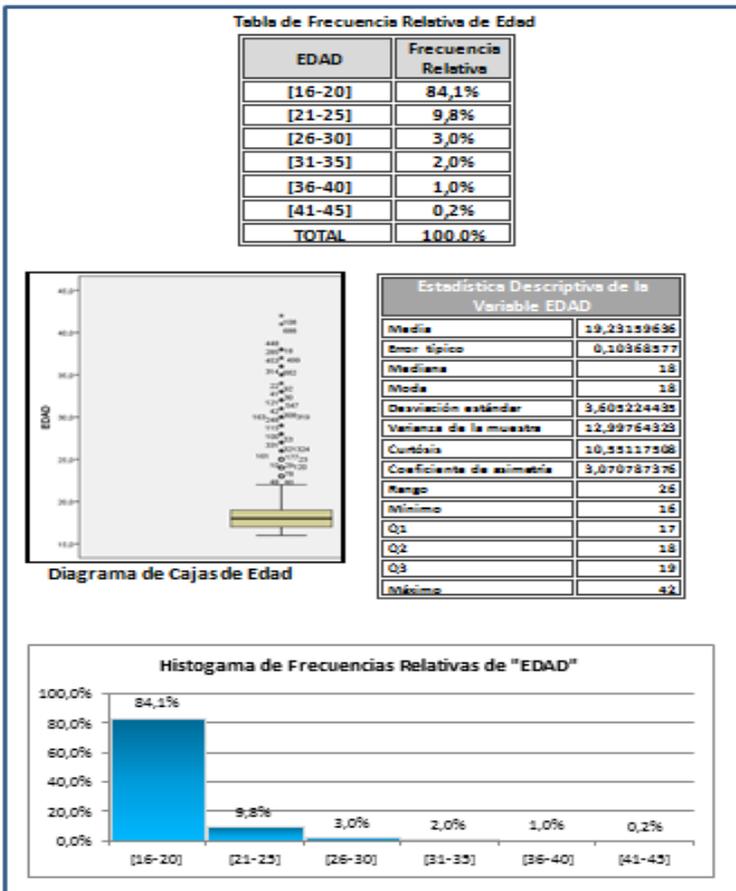


Este tipo de variable es de tipo cualitativa nominal, la cual tiene dos categorías, por lo que podemos decir que es una variable dicotómica. Al observar la tabla de frecuencias (Cuadro 2), de esta variable podemos analizar que, del total de entrevistados (1209 estudiantes), que forman parte de la muestra piloto se puede concluir que 728 estudiantes son de género femenino y que 481 estudiantes encuestados son de género masculino, estos datos son contrastados con el gráfico del histograma de frecuencia, por lo que podemos observar que el porcentaje del género femenino corresponde al 60,20% del total de la muestra y que el porcentaje

del género masculino es del 39,80%, lo que concluimos, referente a la deserción estudiantil de los cursos de nivelación, que la mayoría de las personas encuestadas correspondieron al sexo femenino.

9.8.2.3 Variable: Edad del entrevistado

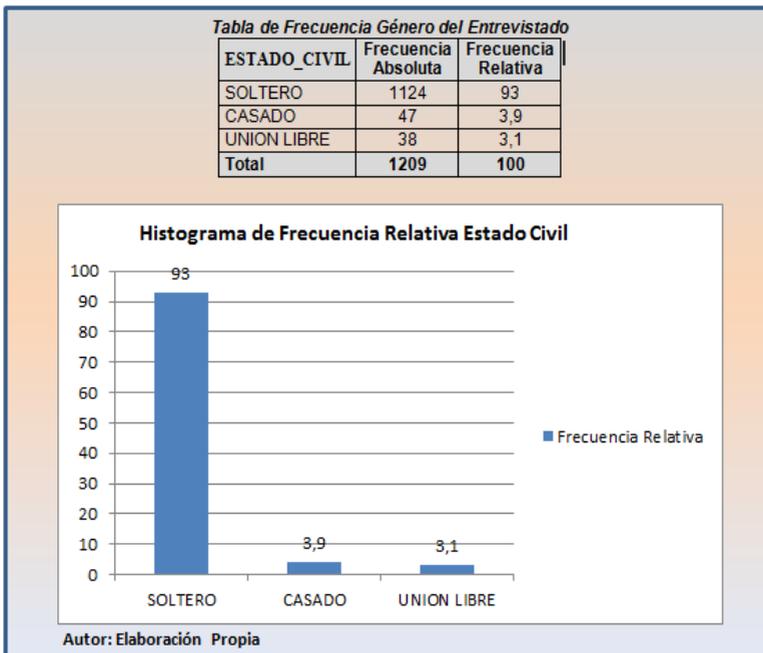
Cuadro 3: Edad del Entrevistado



En este resumen estadístico para la variable EDAD, observamos en el cuadro 4, que de un total de 1209 estudiantes encuestados la edad promedio es de 19.23 años y una mediana de 18 años, si analizamos la media y la mediana, nos damos cuenta que la media es mayor, lo que nos da a interpretar el primer estadístico de forma, que es el coeficiente de asimetría de Fisher, con un valor de 3.071 positivo, lo que quiere decir que la distribución de probabilidad bajo la curva normal, presenta una asimétrica hacia la derecha, esto implica que la mayoría de los datos se encuentran concentrados hacia la izquierda de la distribución es decir la mayor cantidad de personas encuestadas estuvieron entre 16 y 21 años de edad; con respecto a al segundo estadístico de forma (Curtósis), podemos indicar que el grado de concentración que presentan los valores alrededor de la media es mayor a cero (10.55) por lo que la curva de Gauss está bien elevada, es decir tiene una distribución Leptocúrtica, con respecto al valor de la desviación estándar podemos decir que la distancia de cada punto respecto a la media es de 19 ± 3.6 años. Otro tipo de estadístico muy importante para este análisis son los de posición, entre los que tenemos a los cuartiles cuya representación gráfica es el diagrama de cajas (Gráfico #11), por lo que podemos observar que la parte derecha de la caja es mayor que la izquierda; ello quiere decir que las edades comprendidas entre el 50% y el 75% de la población están más dispersas que entre el 25% y el 50%, con respecto a los bigotes podemos notar que el de la izquierda (X_{\min}, Q_1) es más corto que el de la derecha; por ello el 25% de los más jóvenes están más concentrados que el 25% de los mayores, también notamos que existen valores atípicos a la derecha del bigote, es decir valores que se alejan de los datos normales de la muestra

9.8.2.4 Variable: Estado Civil del entrevistado

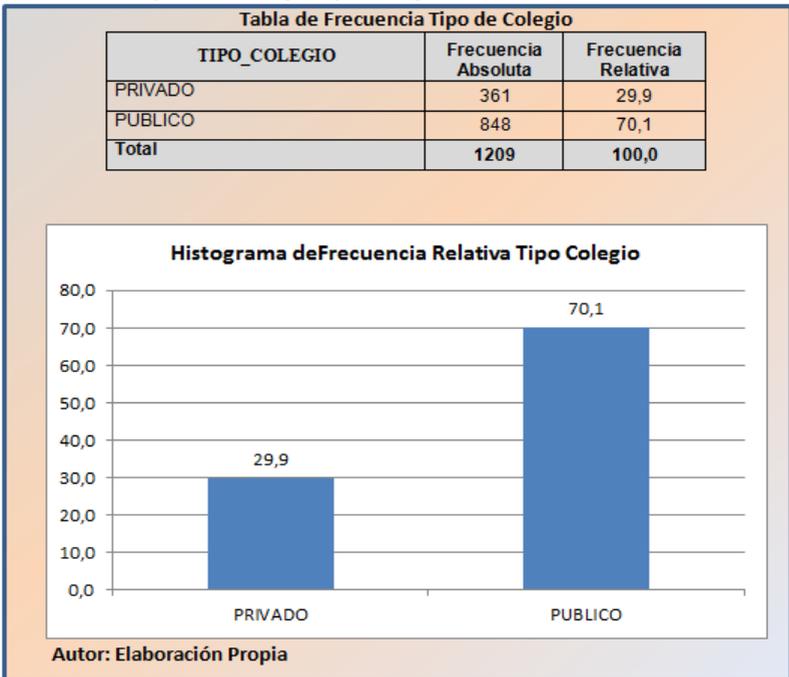
Cuadro 4: Estado Civil del Entrevistado



De acuerdo a los resultados obtenidos en esta variable cualitativa nominal, podemos concluir que el 93% de los entrevistados que respondieron cuál es su estado civil actual, son solteros, mientras que el 3.9% son casados y apenas el 3.1% están en unión libre.

9.8.2.5 Variable: Tipo Colegio

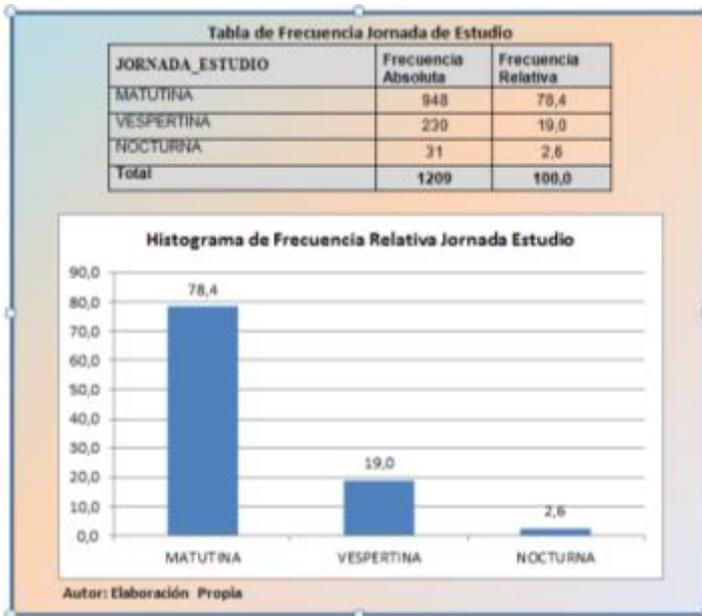
Cuadro 5: Tipo de Colegio que se graduó el Entrevistado



Del total de la muestra de estudiantes encuestados (1209), el 70.1% de los estudiantes provienen de colegios públicos y el 29.9% provienen de colegios privados.

9.8.2.6 Variable: Jornada de Estudio

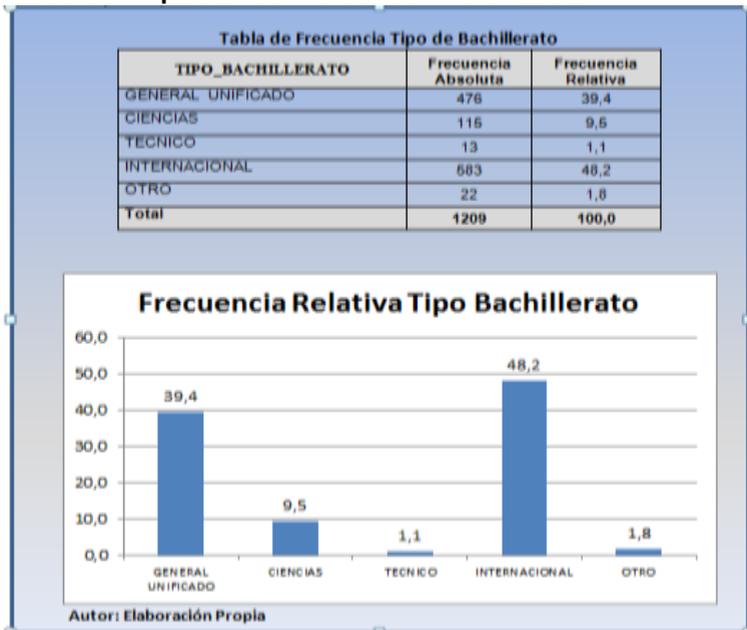
Cuadro 6: Jornada de estudio del Entrevistado



De acuerdo al cuadro 6 se puede observar que cada 8 estudiantes de 10 (78.4%), estudiaron en colegios cuya jornada era matutina, el 19% provienen de jornada vespertina y el 2.6% provienen de jornada Nocturna.

9.8.2.7 Variable: Tipo Bachillerato

Cuadro 7: Tipo de Bachillerato del Entrevistado

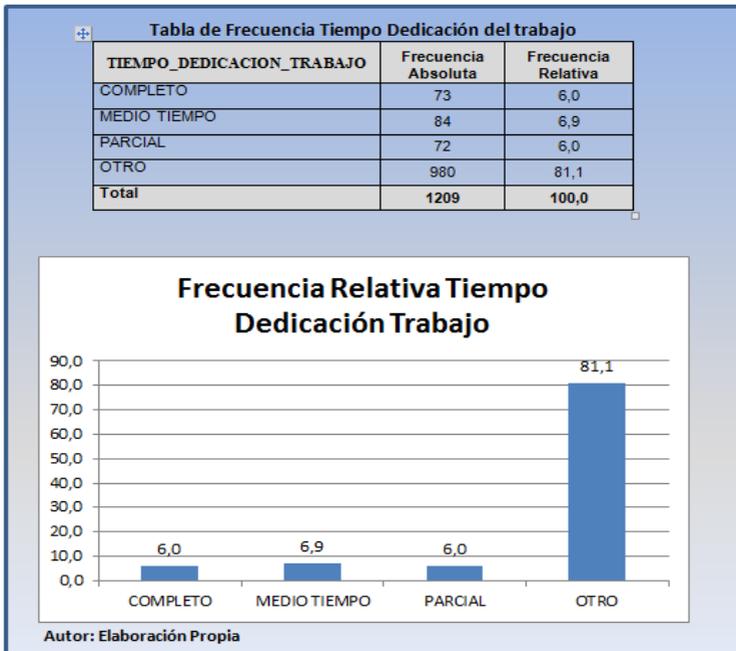


Para el análisis de esta variable, nos podemos dar cuenta, de acuerdo al cuadro 7. El 48.2% de los estudiantes se especializaron en un bachillerato internacional, el 39.4% en un bachillerato general unificado, el 9.5% en un bachillerato en ciencias y 1.1% en un bachillerato técnico.

9.8.3 Sección: INFORMACIÓN FAMILIAR

9.8.3.1 Variable: Tiempo Dedicación Trabajo

Cuadro 8: Tipo de Bachillerato del Entrevistado



Durante el estudio de campo, también nos fue muy importante analizar que pasaba con aquellos estudiantes que de alguna manera desempeñaban alguna tarea laboral, por eso el (cuadro 8), nos indica que de un total de 235 estudiantes que trabajan (19.4%), el 6% trabajaban a tiempo completo, el 6.9% a medio

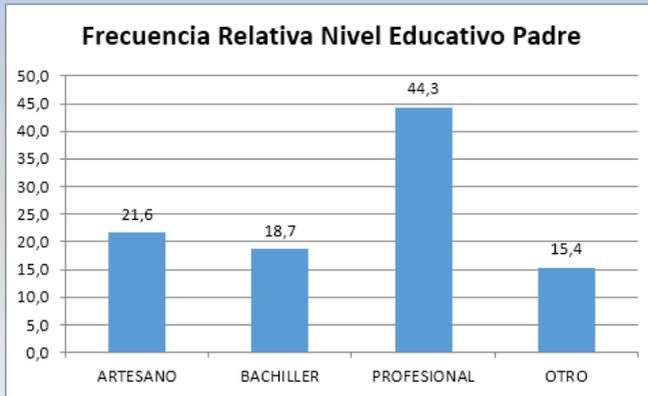
tiempo, y el 81,1% dedica otro tiempo para trabajar menor a 6 horas o simplemente no trabaja.

9.8.3.2 Variable: Nivel Educativo Del Padre

Cuadro 9: Nivel educativo del padre

Tabla de Frecuencia Nivel Educativo del Padre

NIVEL_EDUCATIVO_PADRE	Frecuencia Absoluta	Frecuencia Relativa
ARTESANO	261	21,6
BACHILLER	226	18,7
PROFESIONAL	536	44,3
OTRO	186	15,4
Total	1209	100,0

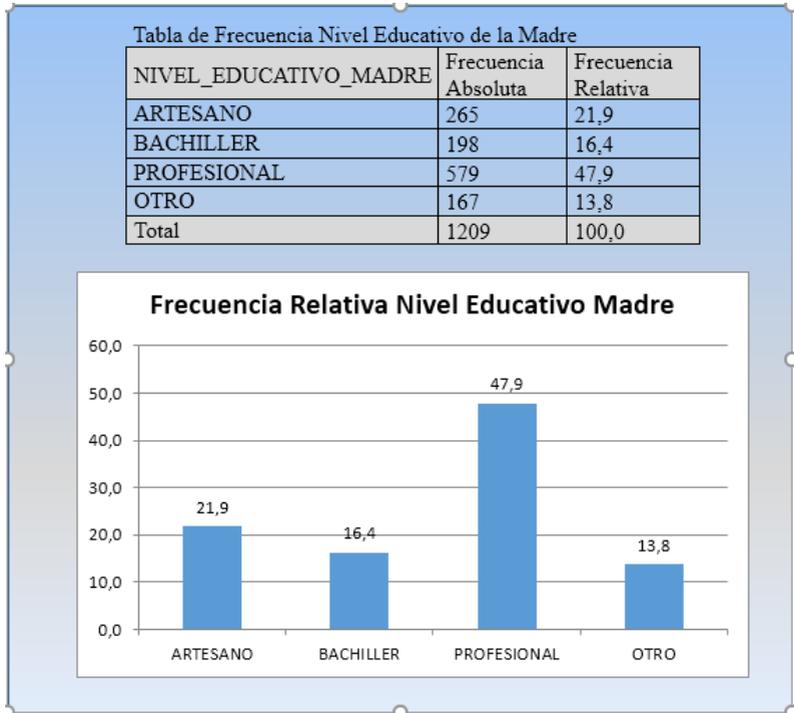


Al analizar la proporción referente al nivel educativo del padre, se conoce que los entrevistados respondieron que el 44,3% tiene un nivel educativo profesional, el 21.6% es artesano el 18.7% son

bachilleres y el 15.4% tienen otro nivel educativo, posiblemente primaria o no tienen un nivel educativo formal.

9.8.3.3 Variable: Nivel Educativo De la Madre

Cuadro 10: Nivel educativo de la madre



Tal como se observa en el cuadro 10, prácticamente 5 de cada 10 madres encuestadas (47.9%), indican que su nivel educativo es profesional, el porcentaje de las madres que manifiestan tener un

Análisis Estadístico Univinario

nivel educativo de artesano es de 21,9%, nivel educativo de bachiller el 16,4% y otro tipo de nivel educativo, es de un 13,8%.

10 ANÁLISIS DE DATOS CUALITATIVOS

Si bien es verdad, el análisis estadístico Univariado para una variable cualitativa no nos proporciona tanta información como lo hace un análisis Univariado para variables cuantitativas, ya que en un análisis cualitativo no le es posible determinar sus estadísticos, debido a que el resultado de un análisis cualitativo es teórico y no numérico como lo es el análisis cuantitativo, en estos casos se ha proporcionado otras formas muy interesantes de cómo hacer un análisis cualitativo, y es a través de la tabla de contingencia.

10.1 CONSTRUCCIÓN DE UNA TABLA DE CONTINGENCIA

La tabla de contingencia es una tabla de doble entrada, donde en cada casilla figurará el número de casos o individuos que poseen un nivel de uno de los factores o características analizadas y otro nivel del otro factor analizado, es importante tener claro que la tabla de contingencia nos ayuda a estudiar dos variables de tipo cualitativa, sean estas dicotómicas o politómicas, por ejemplo tenemos que en las ciencias sociales a menudo se desea determinar la existencia o no de relación entre dos variables, para lo cual se realizan diferentes pruebas de significación estadística. La forma en que se relacionan dos variables se denomina

asociación entre dos variables. El interés de este tipo de análisis se centra principalmente en la forma en que se distribuye la variable dependiente en función de las diferentes categorías valores de la variable independiente, es decir aquella variable que atrae la atención de la persona que hace un análisis estadístico, y la cual quiere dar una interpretación bastante robusta y convincente.

En base al análisis estadístico que realizamos de la tabla de contingencia se puede encontrar un resultado de tipo “Hipotético” o “Teórico”. En el caso hipotético, podemos indicar que; En el estudio de la relación de dos variables cualitativas intentamos encontrar a través de una hipótesis una respuesta tentativa a la pregunta correspondiente al problema, es decir un posible resultado que se espera encontrar, y en el caso teórico nos basamos en el estudio de las relaciones entre dos variable cualitativas, con la finalidad de determinar si las variaciones en los valores de la variable independiente se encuentran o no asociados a las variaciones en los valores de la variable dependiente.

Sin embargo, los resultados que aparecen en este libro, son en base a un análisis estadístico teórico, es decir analizamos la tabla de contingencia tratando de encontrar la relación existente entre un par de variables cualitativas.

10.2 CASO DE ESTUDIO: FACTORES QUE INCIDEN EN EL RENDIMIENTO ACADÉMICO EN LOS ESTUDIANTES DE NIVEL SUPERIOR CON RESPECTO A SU ESTATUS LABORAL

Este caso de estudio, lo analizaremos mediante el uso de la tabla de contingencia, la finalidad es que el lector pueda tener una guía de cómo hacer un análisis estadístico y una buena interpretación de las variables cualitativas.

10.2.1 Definición de las Variables Cualitativas que serán objeto de este estudio.

Jornada de estudio: Indica en que jornada de estudio se encuentra el estudiante de la FCE encuestado que puede ser: MATUTINO, VESPERTINA, NOCTURNA.

Sexo: Indica si el estudiante encuestado es del sexo masculino o femenino.

Estado civil: Indica si el estudiante encuestado es: casado, soltero o unión libre.

Trabaja: Indica si el estudiante encuestado si se encuentra laborando o no.

La cantidad materias debe ser acorde al tiempo disponible: Indica si el estudiante encuestado considera que el número de materias debe ser tomada de acorde al tiempo disponible o no.

El trabajo y la nueva metodología de calificación de la UG son factores que perjudican el rendimiento académico: Indica la opinión del estudiante encuestado sobre si el trabajo y la metodología de calificación de la UG son factores que perjudican el rendimiento académico.

Experiencia laboral tiene relación con la carrera que estudia: indica si el estudiante encuestado se encuentra laborando o si tiene experiencia laboral en relación a la carrera de Ciencias Económicas.

Horas de trabajo para realizar tareas académicas: Nos deja ver si el estudiante encuestado realiza tareas académicas en sus horas de trabajo.

El estrés y el cansancio del trabajo afectan el rendimiento académico: Indica la opinión del estudiante encuestado, si se encuentra totalmente de acuerdo o en desacuerdo sobre si es estrés y el cansancio del trabajo son factores que afectan su rendimiento académico.

Problemas en el ámbito laboral inciden en el bajo rendimiento académico: Nos muestra el nivel de conformidad que tiene el estudiante encuestado con respecto a que los problemas en el ámbito laboral incurren en su bajo rendimiento académico.

Permiso para faltar a jornada laboral para cumplir actividades académicas: Indica si el estudiante encuestado tiene aprobación de permiso por parte de su jefe para faltar a su jornada laboral para cumplir con trabajos (investigaciones, trabajo en grupo, etc.).

El trabajo crea un impacto negativo en el rendimiento académico: Nos permite conocer la opinión del estudiante encuestado sobre si el considera que el trabajo crea un impacto negativo en el rendimiento académico.

10.2.2 La población objetivo son los alumnos de la Facultad de Ciencias Económicas de la Universidad de Guayaquil (FCE).

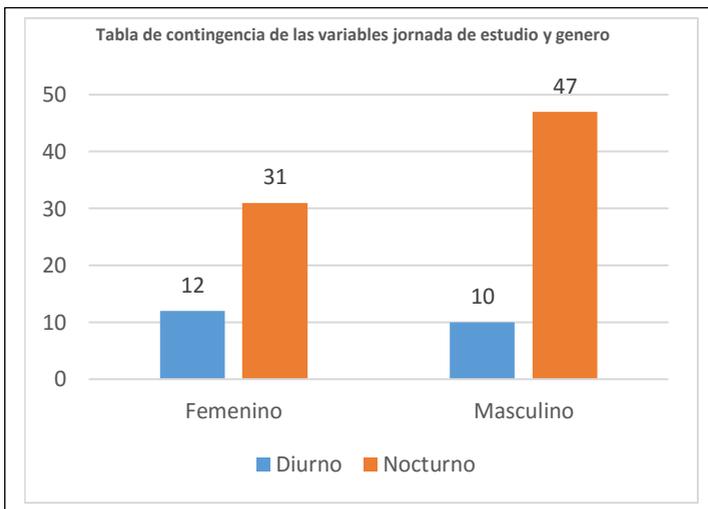
10.2.3 El tamaño de la muestra es de 100 estudiantes

10.2.4 El instrumento de medición a utilizar es la Encuesta.

10.3 ANALISIS DE TABLAS DE CONTINGENCIA

10.3.1 Tabla de contingencia de las variables jornada de estudio y genero

Jornada estudio	Género		Total
	Femenino	Masculino	
Diurno	12	10	22
Nocturno	31	47	78
Total	43	57	100



Identificación de las variables explicativa y de respuesta:

1. La variable explicativa X es el género (F/M).
2. La variable respuesta (Y) es la Jornada de estudio (Diurno/Nocturno).

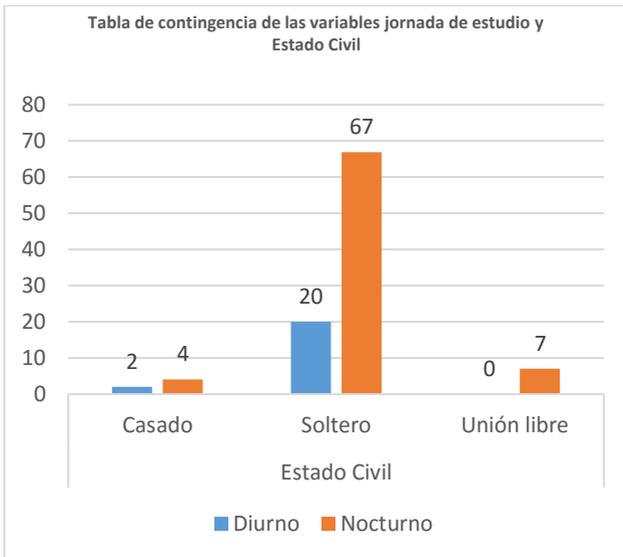
Interpretación de los datos:

En esta tabla se puede observar en primer lugar que de los 100 estudiantes de la FCE de lo que se tiene información, 57 son hombres y 43 son mujeres. Así mismo se sabe que 22 de los estudiantes encuestados, estudian en jornada matutina y 78 en la nocturna. Y que de los 78 estudiantes que pertenecen a la jornada nocturna, 47 de ellos son hombres y 31 son mujeres; mientras que, en el caso de jornada diurna, 10 son hombres y 12 mujeres. Lo que nos permite ver que existe un mayor número de estudiantes que prefieren estudiar en la jornada nocturna y que la mayoría de los estudiantes son del género masculino. La tabla de contingencia nos permite tener información cruzada sobre ambas variables.

10.3.2 Tabla de contingencia de las variables jornada de estudio y Estado Civil

	Estado Civil	Total
--	---------------------	--------------

Jornada estudio	Casado	Soltero	Unión libre	
Diurno	2	20	0	22
Nocturno	4	67	7	78
Total	6	87	7	100



Identificación de las variables explicativa y de respuesta:

1. La variable explicativa X es el Estado Civil (Casado, Soltero, Unión libre).

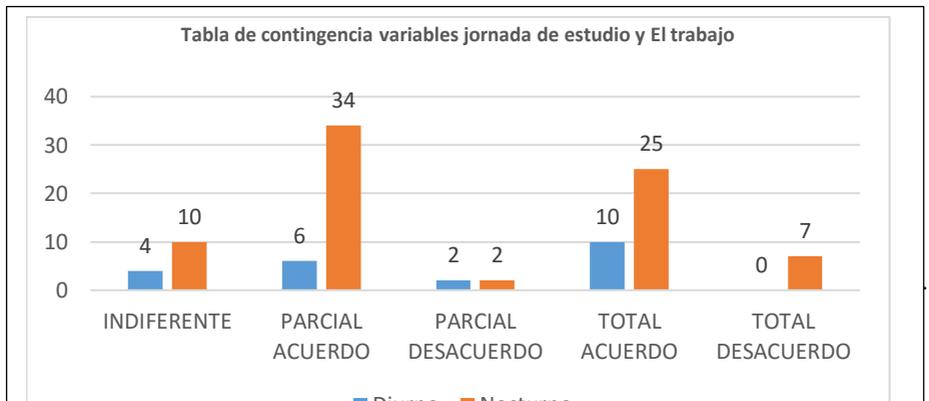
- La variable respuesta (Y) es la jornada de estudio (Diurno/Nocturno).

Interpretación de los datos:

En esta tabla y en grafico de barras se puede observar que de los 100 estudiantes encuestados 78 estudian en la jornada nocturna y 22 en la diurna. La tabla anterior nos permite tener información cruzada sobre ambas variables: de los 78 estudiantes de la jornada nocturna: 6 son solteros, 87 son solteros y 7 están en unión libre, mientras que de los 22 de la jornada diurna: 2 son casados, 20 solteros y 7 en unión libre. De lo que podemos concluir que en ambas jornadas de estudio el mayor número de estudiantes son solteros por lo cual el estado civil no sería un factor que incida en el bajo rendimiento académico de los estudiantes que laboran.

10.3.3 Tabla de contingencia de las variables jornada de estudio y El trabajo y la metodología de calificación son factores que perjudican el rendimiento académico.

Jornada	El trabajo y la metodología					Total
	INDIFERENTE	PARCIAL ACUERDO	PARCIAL DESACUERDO	TOTAL ACUERDO	TOTAL DESACUERDO	
Diurno	4	6	2	10	0	22
Nocturno	10	34	2	25	7	78
Total	14	40	4	35	7	100



Identificación de las variables explicativa y de respuesta:

1. La variable explicativa X trabajo y metodología (indiferente, Parcial Acuerdo, Parcial desacuerdo, total acuerdo, total desacuerdo).
2. La variable respuesta (Y) es la jornada de estudio (Diurno/Nocturno).

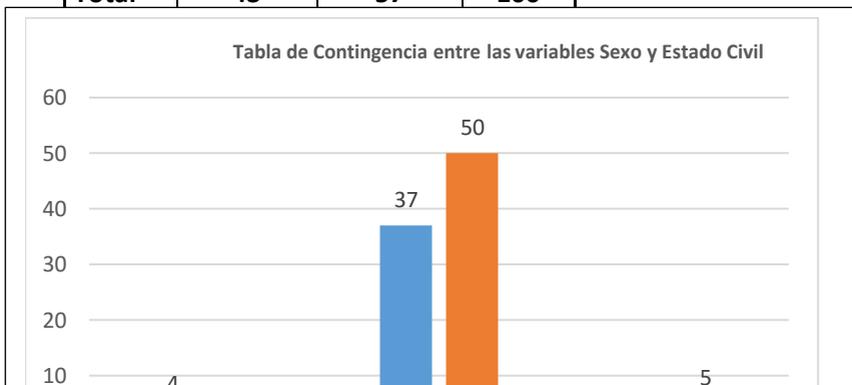
Interpretación de los datos:

En esta tabla y el gráfico de barras, se puede observar que de los 100 estudiantes encuestados 78 estudian en la jornada nocturna y 22 en la diurna. También nos permite saber que 35 de estos estudiantes se encuentran totalmente de acuerdo en considerar que el trabajo y la metodología de estudio son factores que perjudican su rendimiento académico, 40 se encuentran parcialmente acuerdo, 14 lo consideran indiferente, 4 están en parcial desacuerdo y 7 en total desacuerdo. También podemos obtener alguna información cruzada sobre ambas variables: de los 78 estudiantes de la jornada nocturna: 25 están en total acuerdo en considerar que el trabajo y la metodología de estudio son factores que perjudican su rendimiento académico, 34 en parcial acuerdo, 10 lo creen que le es indiferente, 2 están en parcial desacuerdo y 7 están en total desacuerdo, mientras que de los 22 de la jornada diurna: 10 están el total acuerdo, 6 están en parcial acuerdo, 4 lo creen indiferente, 2 en parcial desacuerdo y ninguno

está en total desacuerdo, esto nos da a entender y a interpretar que las personas que estudian en la noche tienen problemas en cuanto se refiere a lo dificultoso que se les hace estudiar y poder cumplir con sus actividades académicas.

10.3.4 Tabla de Contingencia entre las variables Sexo y Estado Civil

Estado Civil	Genero		Total
	Femenino	Masculino	
Casado	4	2	6
Soltero	37	50	87
Unión libre	2	5	7
Total	43	57	100



Identificación de las variables explicativa y de respuesta:

1. La variable explicativa X es el género (F/M).
2. La variable respuesta (Y) es estado civil (casado, soltero, unión libre).

Interpretación de los datos:

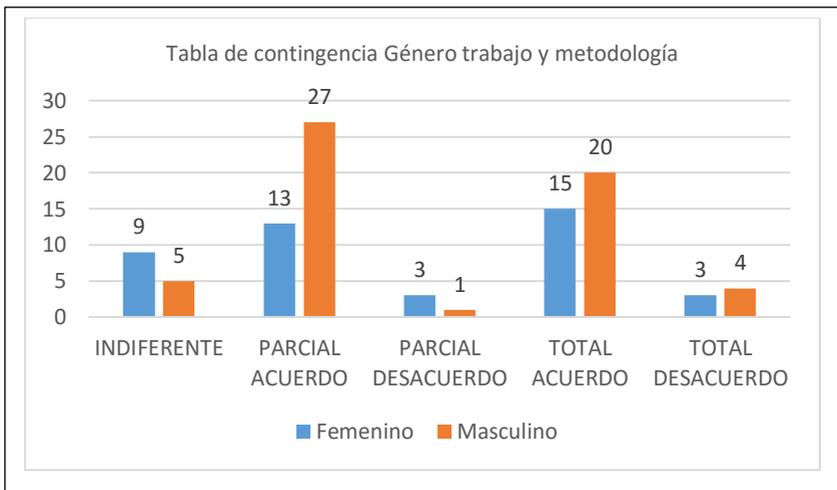
Si observamos en la gráfica y en la tabla de contingencia, se puede apreciar que hay un bajo número de personas del género masculino que están casados y estudian, apenas tenemos 2 estudiantes; en lo que se refiere al género femenino tenemos un total de 4 estudiantes, aunque en realidad es poca la población de las personas que se encuentran casada o comprometidas y que están estudiando, si observamos nuevamente la tabla de contingencia, y nos concentramos en los estudiantes que tienen una relación en unión libre, apenas tenemos 7 estudiantes de los cuales 2 son del sexo femenino y 5 del sexo masculino, teniendo un total entre alumnos casados y en unión libre de apenas 13 personas de un total de 100 personas encuestadas, a lo que corresponde apenas 13% de toda la población.

Ahora analizaremos el número de personas Solteras, de un total de 87 estudiantes solteros, 50 estudiantes son del género

masculino es decir el 57%, y 37 son del género femenino es decir 42.5%.

10.3.5 Comparación entre las variables Genero y El trabajo y la metodología de calificación son factores que perjudican el rendimiento académico

Genero	El trabajo y la metodología					Total
	INDIFERENTE	PARCIAL ACUERDO	PARCIAL DESACUERDO	TOTAL ACUERDO	TOTAL DESACUERDO	
Femenino	9	13	3	15	3	43
Masculino	5	27	1	20	4	57
Total	14	40	4	35	7	100



Identificación de las variables explicativa y de respuesta:

1. La variable explicativa X es el trabajo y la metodología (indiferente, Parcial Acuerdo, Parcial desacuerdo, total acuerdo, total desacuerdo).
2. La variable respuesta (Y) es el género (F/M).

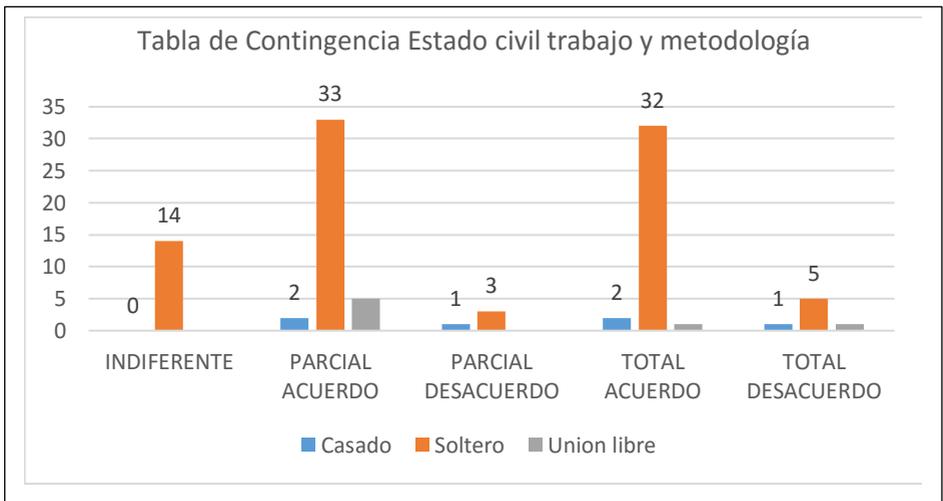
Interpretación de los datos:

Al hacer la comparación entre el sexo y la metodología de calificación, con el grafico y los datos obtenidos en la tabla de contingencia, podemos inferir, en que, de un total de 57 personas de género masculino, el 82% $((27 + 20) / 57)$, están en total acuerdo y en parcial acuerdo con que el trabajo y la metodología de calificación son factores que perjudican el rendimiento académico, así mismo de un total de 43 personas del género femenino, 28 están en total acuerdo y en parcial acuerdo con que el trabajo y la metodología de calificación son factores que perjudican el rendimiento académico.

10.3.6 Comparación entre las variables Estado Civil y El trabajo y la metodología de calificación son factores que perjudican el rendimiento académico.

Análisis Estadístico Univinario

Estado Civil	El trabajo y la metodología de calificación son factores que perjudican el rendimiento académico					Total
	INDIFERENTE	PARCIAL ACUERDO	PARCIAL DESACUERDO	TOTAL ACUERDO	TOTAL DESACUERDO	
Casado	0	2	1	2	1	6
Soltero	14	33	3	32	5	87
Unión libre	0	5	0	1	1	7
Total	14	40	4	35	7	100



Identificación de las variables explicativa y de respuesta:

1. La variable explicativa X es el género (F/M).
2. La variable respuesta (Y) es la jornada de estudio (Diurno/Nocturno).

Interpretación de los datos:

La metodología de calificación ha venido en un proceso de transformación para que tanto el estudiante como el profesor se sientan cómodos. Con relación a las personas casadas, estas consideran que la metodología que se está llevando es factible hacia sus beneficios, en este caso tienen un valor de 2 en Parcial Acuerdo y Total Acuerdo, pero, ya que si lo analizamos mejor es un valor muy pequeño y no tan considerable para este análisis.

Las personas solteras al ser un mayor número, nos damos cuenta que no están conformes con la metodología de calificación ya que su trabajo les limita hacer mayor cantidad de tareas enviadas por el profesor y si verificamos los números decimos que de un total de 87 personas encuestadas y que son solteros, 32 indican un Total acuerdo y con 33 el Parcial acuerdo, de la característica que se está analizando.

Por último, se tienen a las personas que son unión libre, estos consideran que el método de calificación si es viable para continuar estudiando, con un valor de 5 en Total Desacuerdo.

55.CONCLUSIONES

Mediante los conceptos estudiados y aplicados en la entrega de este libro, hemos logrado adquirir nuevos conocimientos de Excel en cuanto a la aplicación de la estadística descriptiva, adquiriendo nuevos instrumentos para el análisis de datos. nos pareció importante hacer uso de herramientas para obtener muestras representativas.

También fue posible realizar el análisis y la interpretación de los datos, logrando de esta manera obtener un buen enfoque conceptual y práctico que nos permitirá resolver situaciones, tanto en nuestra vida personal como profesional.

Mediante el presente trabajo práctico hemos podido reflejar la incorporación de los conceptos teóricos y prácticos enseñados por la cátedra de probabilidad y estadística impartido en el tercer semestre de la carrera de ingeniería en Sistemas computacionales de la universidad de Guayaquil, y comprender de mejor manera cada una de las unidades incluidas en los capítulos de este libro pudiéndose así conocer las metodologías que permiten un adecuado análisis de los datos estadísticos. Por otra parte, el uso de una aplicación de oficina como Excel y sus funciones estadísticas nos sirvió para hacer un estudio completo de casos de estudios, muy útil en todos los campos.

En lo pedagógico y didáctico se genera un ámbito propicio para coordinar, fomentar y apoyar las actividades de enseñanza y aprendizaje de probabilidad y estadística, enriqueciendo el día a día al estudiante universitario.

56. Bibliografía

Anderson, D., Sweeney, D., & Williams, T. (2004). *Estadística para administración y economía*. México: CENGAGE Learning.

Casal, J., & Mateu, E. (2003). Tipos de Muestreo. 3-7.

Cassany, D. (2007). *Afilar el Lapicero: Guía de redacción para los Profesionales*. Barcelona, España: Anagrama.

- Fernandez, L. (2007). ¿Como se elabora un cuestionario? *Butlletí LaRecerca*, 1-9.
- Garcia, G., López, C., & Zamorano, C. (2011). *Estadística Básica para estudiantes de Ciencia*. Madrid.
- Hernandez, R., Fernandez, C., & Baptista, P. (2003). *Metodología de la Investigación* (Tercera ed.). México: McGraw-Hill.
- Lagares, P., & Albandoz, J. (2001). Población y muestra. Técnicas de Muestreos.
- Macassi, S., & Cristina, M. M. (1997). *Cómo elaborar muestras para los sondeos de audiencia*. Quito.
- Martinez, F. (2002). *El cuestionario. Un instrumento para la investigación en las ciencias sociales*. Barcelona: Laertes.
- Martinez, M., & Marí, M. (2010). La distribución Poisson. *Universidad Politécnica de Valencia*.
- Padilla, M. T., & Gil, J. (2007). Análisis de Datos en la Investigación.
- Rabolini, N. M. (2009). Técnicas de muestreo y determinación del.
- Torres, M. (2007). Tamaño de una muestra para una Investigación de mercado.
- Walpole, R., Myers, R., & Myers, S. (1999). *Probabilidad y Estadística para Ingenieros* (Sexta ed.). Mexico: PEARSON.

Walpole, R., Myers, R., Myers, S., & Keying, Y. (2012). *Probabilidad y Estadística para Ingenieros*. México: Pearson.

Webster, A. (2001). *Estadísticas Aplicada a los Negocios y Economía*. 21-29.

Fernández, S. F., Sánchez, J. M. C., Córdoba, A., Cordero, J. M., & Largo, A. C. (2002). *Estadística descriptiva*. ESIC Editorial.

Gómez M., *Inferencia Estadística*, 2005

Delgado R., *Iniciación A La Probabilidad Y Estadística-Primera Edición*, 2004

Evans, M. J., & Rosenthal, J. S. . (2005). *Probabilidad y estadística* (reimpresión ed.). (Reverte, Ed., & X. T. Morer, Trad.)

Cáceres, R. Á. (2007). *Estadística aplicada a las ciencias de la salud*. Díaz de Santos.

R Johnson, R., & Kubby, P. (2008). *Variable Cualitativa*. (C. L. Editores, Ed., & J. H. Muñoz, Trad.) New York.

Hidalgo Guevara, W. A. (2009). *Unidad de investigación* (Hidalgo ed.). Guayaquil: Hidalgo.

Alvarado Valencia, J. A., & Obagi Araújo, J. (2008). *Fundamentos inferencia estadística*. (I. 978-956-716-094-4, Ed.) Pontificia universidad Javeriana.

Muñoz, D. R. (2005). *Manual de estadística*. (J. C. Coll, Ed.)
Obtenido de
<https://books.google.com.ec/books?id=WdgpP8dknR08C>

B Mode, E. (1982). *Elementos de probabilidad y estadística* (262 ed.). Reverté.

Peña Guerrero A, Buitrago V, Paulete M. (2007) *Estadística Básica*, 1ra Edición, Instituto Tecnológico Metropolitano, Colombia Pp. 24.

Acerca de los autores

ING. LORENZO JOVANNY CEVALLOS TORRES MSC.

(Primer Autor)



Máster en Gestión de la productividad y la Calidad, Escuela Superior Politécnica del Litoral - Ecuador. Máster Universitario en Modelado Computacional en Ingeniería. Universidad de Cádiz - España. Ingeniero en Estadística Informática, Escuela Superior Politécnica del Litoral - Ecuador. Docente de la Universidad de Guayaquil, Facultad de Ciencias Matemáticas y Física, Carrera de Ingeniería en Sistemas Computacionales.

ING. NELLY AMERICA VALENCIA MARTINEZ MSC.

(Segundo Autor)



Educación Superior. Magister en Docencia de las Matemáticas, Profesora de Segunda Enseñanza Físico Matemático, Ingeniera en Sistemas Computacionales. Docente de la Universidad de Guayaquil, Facultad de Ciencias Matemáticas y Física, Carrera de Ingeniería en Sistemas Computacionales Coordinadora de Redes y Masa Crítica, Universidad de Guayaquil, Contraparte Oficial Programa Prometeos UG

DR. ROOSEVELT LINDORFO BARROS MORALES MSC

(Tercer Autor)



Doctor en Ciencias de la Educación Especialización Historia y Geografía, Especialista en Proyectos de Desarrollo Educativos y Sociales. Licenciado en Ciencias de la Educación Especialización Administración y Supervisión Educativa, Licenciado en Ciencias de la Educación Especialidad Historia y Geografía Administrado Educativo. Docente Universitario en la Facultad de Filosofía Letras y Ciencias de la Educación, Rector Fundador de la Unidad Educativa Organización de los Estados Norteamericanos y Escuela Guayas.

ISBN: 978-9978-59-143-7



VICERRECTORADO DE
INVESTIGACIÓN
GESTIÓN DEL
CONOCIMIENTO
Y POSGRADO

