

**FACTORES QUE INCIDEN EN EL MAL USO DE LA INFORMACIÓN EN TRABAJOS DE INVESTIGACIÓN CIENTÍFICA**

EL MAL USO DE LA INFORMACIÓN EN TRABAJOS DE INVESTIGACIÓN CIENTÍFICA

AUTORES: Lorenzo Cevallos Torres<sup>1</sup>Alfonso Guijarro Rodríguez<sup>2</sup>Leili López Domínguez Rivas<sup>3</sup>DIRECCIÓN PARA CORRESPONDENCIA: [lorenzo.cevallos@ug.edu.ec](mailto:lorenzo.cevallos@ug.edu.ec)

Fecha de recepción: 10 - 01 - 2016

Fecha de aceptación: 19 - 02 - 2016

## RESUMEN

Con la finalidad de prevenir el plagio en trabajos de investigación académica, se trata de analizar en este artículo los posibles factores que inciden en que el estudiante haga mal uso de la información en sus tareas universitarias. Para ello, se ha considerado necesario realizar un análisis estadístico, con la finalidad de determinar la probabilidad de cometer fraude. Para dar respuesta a lo planteado se diseñó una investigación, cuya población estuvo determinada por los estudiantes de la universidad de Guayaquil y la muestra estuvo compuesta por un total de 200 alumnos a los cuales se les aplicó una encuesta. Los resultados de la investigación evidenciaron las prácticas deshonestas vinculadas a la elaboración de trabajos académicos, y se comprobó que el uso de internet y la falta de análisis y lectura de los textos científicos, se ha convertido en la fuente principal de los alumnos al momento de plagiar sus trabajos académicos.

**PALABRAS CLAVE:** Citas bibliográficas; infracciones estudiantiles; plagio académico; investigación.

**FACTORS AFFECTING THE MISUSE OF INFORMATION IN SCIENTIFIC RESEARCH WORKS**

## ABSTRACT

In order to prevent plagiarism in academic research, it is analyzed in this paper the possible factors affecting the student misuses the information in their academic tasks. To do this, it has been considered a statistical analysis in order to determine the likelihood of fraud. To address the issues raised an

---

<sup>1</sup> Universidad de Guayaquil. Facultad de Ciencias Matemáticas y Físicas, Carrera de Ingeniería en Sistemas Computacionales. Guayaquil. Ecuador.

<sup>2</sup> Universidad de Guayaquil. Facultad de Ciencias Matemáticas y Físicas, Carrera de Ingeniería en Sistemas Computacionales. Guayaquil. Ecuador. E-mail: [alfonso.guijarror@ug.edu.ec](mailto:alfonso.guijarror@ug.edu.ec)

<sup>3</sup> Universidad de Guayaquil. Facultad de Ciencias Matemáticas y Físicas, Carrera de Ingeniería en Sistemas Computacionales. Guayaquil. Ecuador. E-mail: [lieli.lopezdominguezr@ug.edu.ec](mailto:lieli.lopezdominguezr@ug.edu.ec)

investigation was designed, whose target population consists of students from the University of Guayaquil, through a survey. The sample consisted of a total of 200 students. From the results of the investigation it was noted in relation to dishonest practices linked to the development of academic papers, it is found that the use of internet and lack of reading scientific texts and analysis, has become the main source of when students plagiarize their academic work.

**KEYWORDS:** Bibliographic citations; student breaches; academic plagiarism; investigations.

## INTRODUCCIÓN

Las irregularidades éticas así como el plagio académico en el proceso de investigación es un tema muy común durante el ciclo universitario, los alumnos no se esfuerzan por investigar, dando como resultado copias idénticas de otros investigadores. Como sostiene Kibler (1993) "Una de las principales dificultades a la hora de encarar el tema de la deshonestidad académica es la falta de una definición unívoca y clara del concepto ya que se trata de un constructo basado en principios ético-morales y por consiguiente, asociado a una época, una cultura y una sociedad determinada". A partir de las reflexiones anteriores se busca en este trabajo encontrar los factores que inciden que el estudiante universitario haga mal uso de la información académica, con la finalidad de brindar alternativas para que no cometan fraude y el estudiante pueda realizar trabajos honestos.

La calidad de la educación superior a nivel mundial exige que las universidades del siglo XXI generen conocimiento científico a través de publicaciones, es por esa razón que la universidad de Guayaquil no debe estar al margen de esta situación. Al respecto Hativa. N. (2000) señala "Independientemente del tipo de disciplina o especialidad, las universidades deben formar en sus estudiantes un pensamiento de alto nivel y convertirlos en aprendices autónomos.". Por tanto, los estudiantes universitarios deben crear sus propias habilidades para investigar, organizar, y usar la información que les ofrecen conocimientos, de esta manera se propicia el aprendizaje autónomo.

Las tecnologías de información son una de las herramientas más importantes para los estudiantes universitarios, según estudios realizados la más utilizada es el internet por ser la fuente principal de información, y servir de ayuda para tener un soporte de cómo se redacta un artículo científico y al mismo tiempo adquirir el conocimiento que se ha acumulado en todo el mundo.

Sánchez, J (2000) precisa que "La tecnología es una herramienta con gran capacidad, que al utilizarla con una metodología y diseño adecuado puede ser un buen medio con el cual construir y crear". Es por ello, que para investigar debemos tener claro que la información que se encuentra no nos pertenece y por ello no debemos darle mal uso, y al hablar de mal uso esto implica no plagiar ni crear copias idénticas sobre dicha investigación.

El autor citado anteriormente destaca que “El plagio constituye una acción en la que quién copia ilícitamente se atribuye la paternidad de la obra”, al respecto Barrón, A. (2010) precisa que “Plagiar significa incluir en un documento fragmentos de textos escritos por alguna otra persona sin darle el crédito adecuado”. De igual manera, Huamaní et al., (2008) apunta que “La información necesaria en la realización de investigaciones científicas, cuando son mal empleadas, tributa negativamente a este propósito, sustituyendo en ocasiones el análisis, síntesis, abstracción y generalización del pensamiento científico.”

Por otra parte, Buitrago, J (2004) afirma que “Dado que los recursos electrónicos disponibles son de mayor acceso y con una mayor temática y no existen sanciones determinadas para aquellos que incurran en estas prácticas, se estima que el plagio, o ‘copiar y pegar’ de fuentes electrónicas sea frecuente.”

De allí que en este trabajo se pretende estimar la probabilidad de que un estudiante universitario cometa plagio académico y se pueda pronosticar los factores que inciden a que el mismo incurra en el plagio de información, y así poder hacer énfasis en la necesidad de combatir este mal hábito.

## DESARROLLO

### *Muestra y población*

Para el desarrollo de este proyecto se realizó una encuesta, donde se definió como población a los estudiantes activos de la Universidad de Guayaquil, con un tamaño de población de 1,590 individuos matriculados en las carreras de Ingeniería en Sistemas computacionales, Obstetricia, Educación básica, Publicidad & marketing, durante el curso 2014-2015. El tamaño de la muestra con el que se ha trabajado supone un error estándar del 5%, estimado para un nivel de confianza del 95% y bajo la condición más desfavorable de  $P=Q=0.05$ . La recogida de los datos se realizó en el campus de la Universidad de Guayaquil durante el mes Septiembre del 2014 mediante la aplicación individual y anónima. En la tabla 1 se presenta la distribución de estudiantes tomadas de cada una de las carreras antes mencionadas.

Tabla 1: Distribución de estudiantes por Carrera

CARRERA	CANTIDAD ESTUDIANTES	DE NÚMERO CURSOS	DE
Carrera de Ingeniería en Sistemas	722	13	
Carrera de Obstetricia	196	5	
Carrera de Publicidad & Marketing	248	7	
Carrera de Educación Básica	424	8	
Población objetivo	1590		

Para conocer la cantidad de estudiantes que se deben encuestar procedimos a realizar el cálculo del tamaño de la muestra para poblaciones finitas (cuando se conoce el total de unidades de observación que la integran). Hernández, R. (2008) “La muestra se dirige a un grupo de personas, eventos, sucesos, comunidades entre otras, sobre la cual se ha de recolectar los datos, sin que necesariamente sea representativo del universo o población que se estudia”.

Formula:

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

Como resultado de la fórmula el tamaño de la muestra, corresponde a un total de 200 estudiantes para las cuatro carreras. Para determinar la proporción de estudiantes a encuestar en cada una de las carreras, se determinó considerar a cada carrera como un estrato, lo que aplica un muestreo aleatorio Estratificado. Rubin L. (2004), “Una población se divide en subgrupos denominados estratos y se selecciona al azar una muestra de cada estrato”.

Para el cálculo del muestreo Estratificado se definió:

Dada una población  $U = (y_1, \dots, y_N)$ , se entiende por estratificación una subdivisión de  $U$  en  $L$ , donde  $L$  es el tamaño de cada estrato,

$$U_1 = (y_{11}, \dots, y_{1N_1}) = (y_k / K \in U_1), \dots, U_L = (y_{L1}, \dots, y_{LN_L}) = (y_k / K \in U_L),$$

Verificándose que:

$$N_1 + \dots + N_L = N \quad (\text{Población total})$$

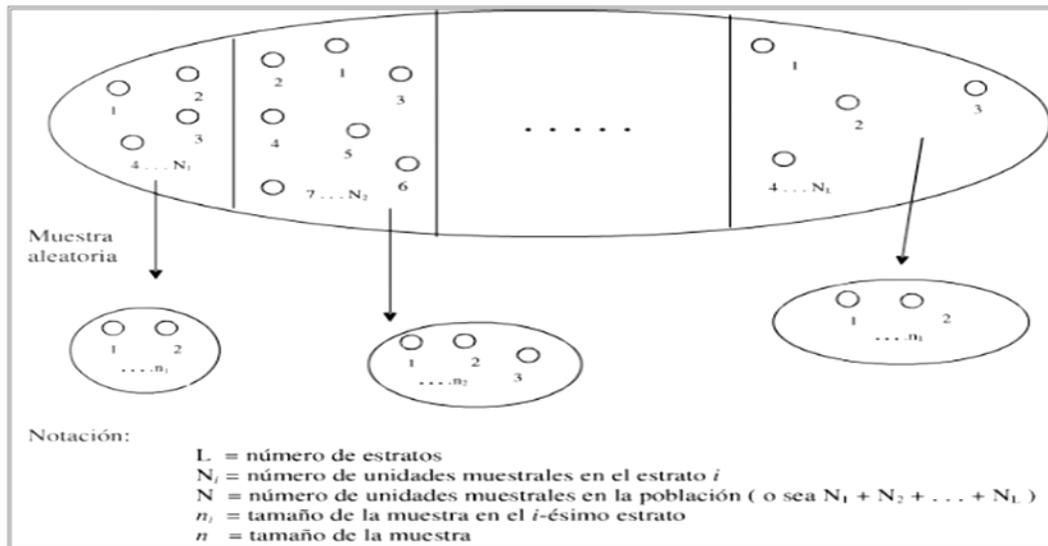
$$n = n_1 \cup \dots \cup n_L \quad (\text{Muestra en el estrato } i\text{-ésimo}).$$

$$n_i = n \left( \frac{N_i}{N} \right) \quad (\text{Tamaño relativo de la muestra en el estrato}).$$

En la gráfica 1, se observa como se hace la distribución de cada uno de los estratos a partir de una población definida.

Se obtiene el tamaño de cada estrato, el mismo que resultado distribuido de la siguiente manera: 91 estudiantes son tomados de la carrera de Ingeniería en Sistemas Computacionales, 25 estudiantes la carrera de Obstetricia, 31 estudiantes de la carrera de Publicidad & Marketing, y 53 estudiantes de Educación Básica. Khan, M (2005) “ No solo ha de tener en cuenta la distribución aleatoria y proporcional del número de muestra sino que tiene también como finalidad que dicha distribución se haga según criterios de contenido social, es decir, criterios que toman en cuenta las proporciones o distribución de los valores de algunas variables (muchas o pocas) en el territorio”.

Gráfica 1: Distribución de la Población en Estratos



### Fuentes de datos y variables de estudio

Los datos para la elaboración del presente artículo provienen de los “Factores que inciden en que los estudiantes de la Universidad de Guayaquil hagan mal uso de la información en sus trabajos de investigación académica” cuya principal finalidad consistía en analizar qué factores tienen mayor incidencia sobre la ejecución de actos académicos deshonestos.

Se estableció para el análisis correspondiente, las siguientes variables que forman parte de nuestro estudio.

- Considera que el método copiar y pegar es el adecuado al presentar un trabajo de investigación académica.
- El uso de fuentes bibliográficas que utiliza para investigar.
- Motivos por los que realiza un trabajo de investigación.
- Horas dedicadas a realizar un trabajo de investigación.
- Tipo de fuentes bibliográficas confiables para hacer un trabajo de investigación.
- Tipo de bases científicas para hacer sus trabajos de investigación.
- Utilización del tiempo necesario para leer y analizar la información que se busca.
- Adquirir conocimiento adecuado al finalizar el trabajo de investigación.
- Utilizar citas bibliográficas en su trabajo de investigación académica.

Con el objetivo de medir la asociación e interacción entre las variables que forman parte de nuestro estudio estadístico, se tomaron en consideración la relación de dichas variables con diversas categorías y rasgos descriptivos del alumnado participante: Género, Facultad perteneciente. (Ver Tabla 2).

Tabla 2: Características y distribución de la muestra del estudio (tomado de la muestra = 200 alumnos)

Variable	Frecuencia	Porcentaje
<b>Sexo:</b>		
Femenino	112	56%
Masculino	88	44%
<b>Carrera</b>		
Ing. Sistemas	91	45%
Obstetricia	25	12%
Publicidad & Marketing	31	16%
Educación Básica	53	27%

### *Análisis de datos*

Respecto a cada una de las variables cualitativas-categóricas se ha efectuado el cálculo de la frecuencia absoluta y la frecuencia relativa, mientras que en las variables cuantitativas de escala se ha establecido el cálculo de los estadísticos de centralización, dispersión, de posición y forma. Para efecto de poder establecer la posible asociación entre las variables en base a las que se operativizan las competencias estudiadas y las características del alumnado o variables categóricas (Tabla 2), se han diseñado tablas de contingencia para cada una de las variables y se ha efectuado la prueba de chi-cuadrado calculándose el coeficiente de correlación de Pearson para medir la asociación lineal entre las variables de estudio.

Para predecir la probabilidad de ocurrencia del evento que es objeto de nuestro estudio (plagio/ no plagio), se aplicó un análisis de regresión logística. Londoño, F. (1995) “La regresión logística es un modelo que se usa comúnmente cuando las variables independientes incluyen medidas tanto continuas como categóricas y la variable de respuesta es categórica como por ejemplo (verdadero/falso). Este modelo estima la probabilidad de que los resultados se presenten en una función exponencial de variables independientes”. Para todos los análisis descritos se he empleado el paquete estadístico “Statistical Package for the Social Sciences” (SPSS versión 22.0)

### *Análisis Estadístico Univariado de la variable edad*

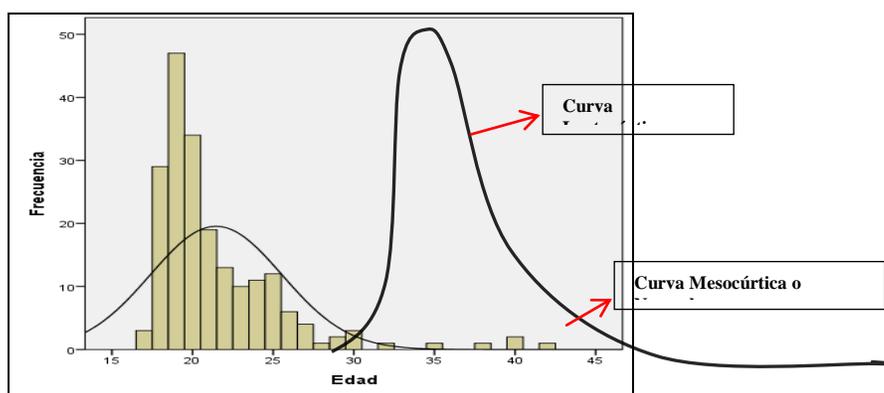
En este resumen estadístico para la variable *EDAD*, se observa en la tabla 3, que de un total de 200 estudiantes encuestados la edad promedio es de 21.46 años y una mediana de 20 años, si se analiza la media y la mediana, nos damos cuenta que la media es mayor, lo que nos da a interpretar el primer

estadístico de forma, que es el coeficiente de asimetría de Fisher, con un valor de 2.41 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 18 y 21 años de edad (gráfico 3); con respecto 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 (7.68) por lo que la curva de Gauss está bien elevada, es decir, tiene una distribución Leptocúrtica (gráfico 2), con respecto al valor de la desviación estándar podemos decir, que la distancia de cada punto respecto a la media es de  $21 \pm 4$  años.

Tabla 3: Estadística Descriptiva Variable EDAD

Estadística Descriptiva Variable EDAD		
N		200
Media		21,46
Mediana		20,00
Moda		19
Desviación estándar		4,082
Varianza		16,662
Coeficiente de Asimetría		2,412
Curtósis		7,686
Q <sub>1</sub>	25	19,00
Q <sub>2</sub>	50	20,00
Q <sub>3</sub>	75	23,00

Gráfico 2: Histograma variable EDAD



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, por lo que podemos observar del gráfico 3, que la parte derecha de la

caja es mayor que la izquierda; lo que 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, se observa que existen valores atípicos a la derecha del bigote, es decir, valores que se alejan de los datos normales de la muestra.

Gráfico 3: Diagrama de cajas EDAD

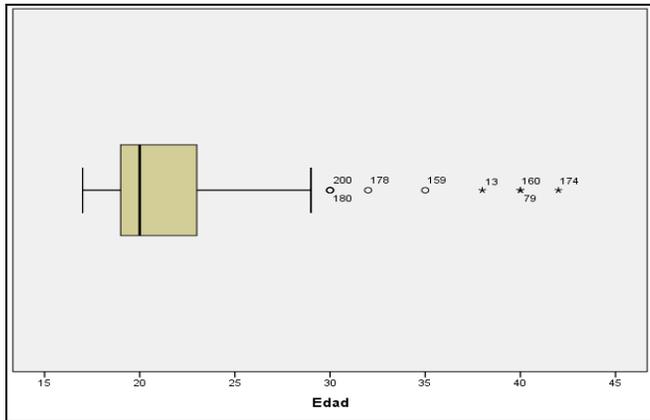


Tabla 4: Frecuencia EDAD

Edad	Frecuencia	Porcentaje
17	3	1,5
18	29	14,5
19	47	23,5
20	34	17,0
21	19	9,5
22	13	6,5
23	10	5,0
24	11	5,5
25	12	6,0
26	6	3,0
27	4	2,0
28	1	,5
29	2	1,0
30	3	1,5
32	1	,5
35	1	,5
38	1	,5
40	2	1,0
42	1	,5
Total	200	100,0

*Análisis de la variable plagio académico por parte del alumno universitario a la hora de elaborar y presentar sus trabajos de investigación.*

Como se muestra en la tabla 5, prácticamente seis de cada diez estudiantes encuestados (61%), admiten que el método de copiar y pegar es el adecuado al presentar sus trabajos académicos, es decir ellos admiten haber copiado fragmento de textos o textos completos y haberlos presentado como propios.

Tabla 5: Frecuencia de plagio en la elaboración de trabajos académicos

¿Usted considera que el método copiar y pegar es el adecuado al presentar un trabajo de investigación académica?	n	Porcentaje
Si	78	39%
No	122	61%
Total	200	100%

*Análisis de las variables que son los factores de incidencia para que el alumno universitario cometa Plagio al momento de elaborar y presentar sus trabajos académicos.*

Como puede comprobarse en la Tabla 6, los estudiantes utilizan el internet como fuente de mayor consulta para hacer sus trabajos de investigación (53.5%), seguido de la utilización de los apuntes que el profesor les da en clase (36%), y pocos son los alumno que consultan revistas científicas (5%) y libros (5.5%).

Tabla 6: Frecuencia de utilización de fuentes bibliográficas en la elaboración de trabajos académicos

¿Qué tipo de fuentes bibliográficas utiliza para realizar sus trabajos de Investigación académica?	n	Porcentaje
Apuntes del Profesor	72	36%
Revistas científicas	10	5%
Internet	107	53,5%
Libros	11	5,5%
Total	200	100

Cerca del 94% de la muestra (93.5%) admite no haber utilizado citas bibliográficas en sus trabajos de investigación académica, y casi uno de cada diez universitarios afirman haber utilizado en sus trabajos investigativos, citas bibliográficas (Tabla 7).

Tabla 7: Frecuencia de utilización de Citas bibliográficas en la elaboración de trabajos académicos

¿Usted hace uso de citas bibliográficas en su trabajo de investigación académica?	n	Porcentaje
Si	13	6,5
No	187	93,5
Total	200	100

De las categorías de la muestra examinada se destaca la existencia de diferencias estadísticamente significativas de las variables Bases de consulta, leer y analizar, y adquirir conocimiento, cuando se asocia con la variable fuentes a utilizar (Ver tabla 8)

Tabla 8: Valor de la prueba de Chi-cuadrado de Pearson para la asociación entre Fuentes a utilizar en trabajos académicos y las variables Bases a consultar.

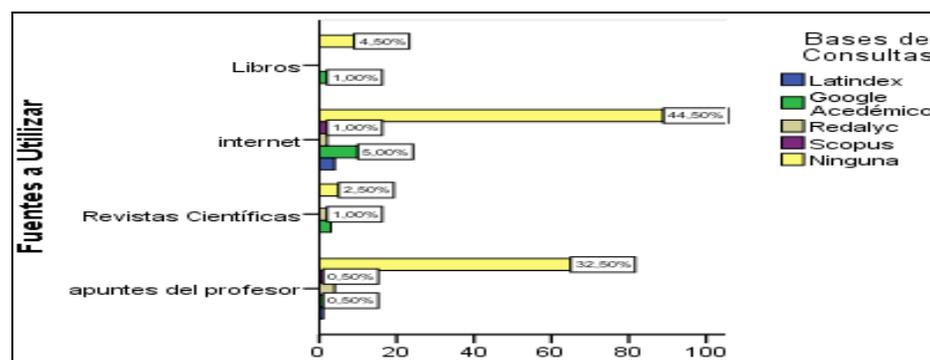
Características de la Muestra	fuentes a utilizar			
	X <sup>2</sup>	p	g.l.	n
Bases de Consulta	23,825*	0,021	12	200
Leer y Analizar	26,806*	0,000	3	200
Adquirir Conocimientos	28,013*	0,006	12	200

Nota: Leer-analizar, y Adquirir Conocimientos.

X<sup>2</sup> Chi-cuadrado de Pearson // p significación // g.l. Grados de libertad // \* Significativa al nivel 0.01 (bilateral) // n número de alumnos de la encuesta

Como puede comprobarse en el Gráfico 4, los estudiantes universitarios cuando realizan un trabajo de investigación académica utilizan el internet como mayor fuente de consultas pero 44.5% no consulta ninguna base científica al momento de hacer sus trabajos, lo mismo sucede cuando toman los apuntes del profesor, el 32.5% de los estudiantes no los compara con ningún tipo de base científica.

Gráfico 4: Gráfico de barras de la Tabla de Contingencia de la variable Fuentes a utilizar Vs. Bases de consultas por parte del estudiante



Según se desprende de los datos obtenidos por el estudio (Ver Gráfico 5), los alumnos de la Universidad de Guayaquil, no leen ni analizan los trabajos de investigación, lo que los conlleva a utilizar una vía más rápida para solo presentar un trabajo, es decir 45.5% utiliza el internet como fuente principal de obtener información necesaria para cometer plagio académico.

Los alumnos que utilizan internet, el 36.5% algunas veces adquieren conocimiento de lo que están investigando, si se observa en el gráfico 4; cuando

el alumno utiliza como fuente de consulta los libros el 1% adquieren un conocimiento de lo que investigan.

Gráfico5: Gráfico de barras de la Tabla de Contingencia de la variable Fuentes a utilizar Vs Leer y analizar

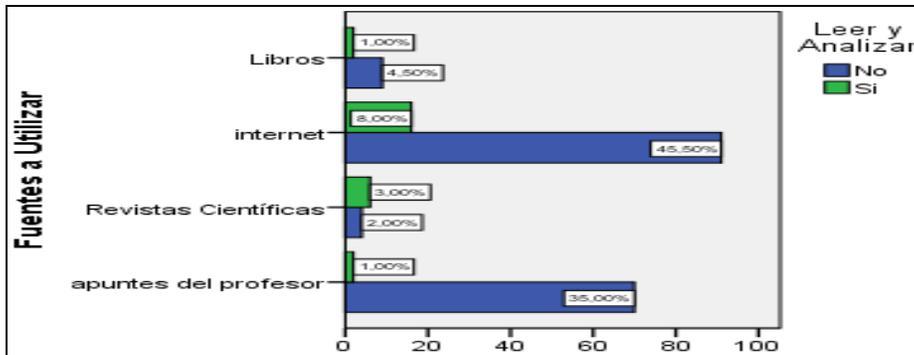
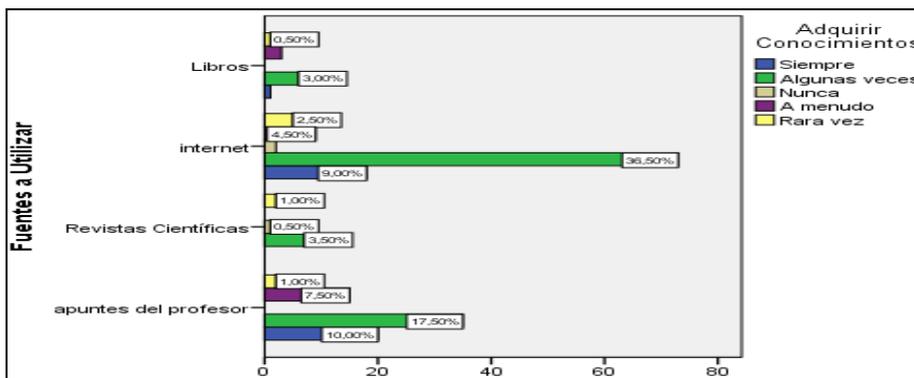


Gráfico 6: Gráfico de barras de la Tabla de Contingencia de la variable Fuentes a utilizar Vs Adquirir conocimientos



Otra categoría importante analizar, se destaca la variable Leer y Analizar, ya que esta variable presenta resultados estadísticamente significativos con las variables Fuentes a utilizar, Bases a consultar y adquirir conocimientos.

Tabla 3: Valor de la prueba de Chi-cuadrado de Pearson para la asociación entre Leer y analizar y las variables Fuentes a utilizar, Bases a consultar, y Adquirir Conocimientos.

Características de la Muestra	Leer y analizar				
	de Chi-square	P	gl.	n	
Fuentes a Utilizar	26,806	0,000	3	200	
Bases de Consulta	118,95	0,000	4	200	
Adquirir Conocimientos	34,79	0,000	4	200	

Nota:  $X^2$  Chi-cuadrado de Pearson // p significación // g.l. Grados de libertad // \* Significativa al nivel 0.01 (bilateral) // n número de alumnos de la encuesta

### Análisis de Correlación de Pearson, de las variables objeto de estudio

Para determinar la relación que existe entre las mismas, se aplicó un análisis de correlación múltiple de Pearson el cual consiste en generar un coeficiente que mida el índice de covariación entre una a más variables relacionadas linealmente y sean indispensables entre sí. Batanero C. (2008) "La correlación o grado de asociación de dos variables se mide utilizando el coeficiente de correlación de Pearson, este coeficiente mide el grado de asociación lineal entre dos variables. Su valor fluctúa en el intervalo [-1, 1]. Cortés, E. (2010). "El coeficiente de Pearson mide la probabilidad de establecer una ecuación lineal entre dos variables, en la que por cada cambio de unidad en una de ellas se espera un cambio de unidad (correlativo) en la otra".

Fórmulas:

$$r_{xy} = \frac{\sum Z_x Z_y}{N}$$

$$r_{xy} = \frac{\sum Z_x Z_y}{N} = \frac{\sum Z_x Z_x}{N} = \frac{\sum Z_x^2}{N} = 1$$

$$r_{xy} = \frac{\sum Z_x Z_y}{N} = 0$$

Respecto de los resultados del análisis de correlación entre las siete variable (Ver tabla 8) que se aplicaron a los 200 alumnos de la Universidad de Guayaquil (UG), se encontraron correlacionadas las variables leer y analizar Vs. Bases de consulta a un nivel de confianza de 0.01 y con un índice de correlación de -0.547 y un valor p de 0.000, esto nos da a entender que el alumnos al no consultar bases científicas y confiables y al no leer ni analizar la información que le proporcionan los textos de estas bases científicas, aumentará la probabilidad de seguir plagiando sus trabajos de investigación académica.

Tabla 4: Tabla de Correlación de las variables objeto de estudio

		Copiar y Pegar	Fuentes a Utilizar	Fuentes de Información	Bases de Consultas	Leer y Analizar	Adquirir Conocimientos	Uso de Citas
Copiar y Pegar	Correlación de Pearson	1	-,099	,040	,059	-,148*	,036	-,039
	Sig. (bilateral)		,165	,578	,405	,036	,616	,587
	N	200	200	200	200	200	200	200
Fuentes a Utilizar	Correlación de Pearson	-,099	1	-,099	-,109	,155*	,015	,066

	Sig. (bilateral)	,165		,162	,126	,029	,830	,355
	N	200	200	200	200	200	200	200
Fuentes de Información	Correlación de Pearson	,040	-,099	1	,134	-,168*	-,129	,077
	Sig. (bilateral)	,578	,162		,058	,018	,069	,280
	N	200	200	200	200	200	200	200
Bases de Consultas	Correlación de Pearson	,059	-,109	,134	1	-,547**	-,027	-,122
	Sig. (bilateral)	,405	,126	,058		,000	,702	,084
	N	200	200	200	200	200	200	200
Leer y Analizar	Correlación de Pearson	-,148*	,155*	-,168*	-,547**	1	,150*	,079
	Sig. (bilateral)	,036	,029	,018	,000		,034	,266
	N	200	200	200	200	200	200	200
Adquirir Conocimientos	Correlación de Pearson	,036	,015	-,129	-,027	,150*	1	,055
	Sig. (bilateral)	,616	,830	,069	,702	,034		,442
	N	200	200	200	200	200	200	200
Uso de Citas	Correlación de Pearson	-,039	,066	,077	-,122	,079	,055	1
	Sig. (bilateral)	,587	,355	,280	,084	,266	,442	
	N	200	200	200	200	200	200	200

\*\* La correlación es significativa en el nivel 0,01 (2 colas).

### *Análisis de los factores que inciden para que el alumno cometa plagio académico*

Para predecir la probabilidad de ocurrencia de un evento, se aplicó un análisis de regresión logística, la cual es una herramienta estadística de análisis Bivariado o multivariado, que permite estimar la probabilidad de que ocurra un evento, Londoño, F. (1995) "La regresión logística es un modelo que se usa

comúnmente cuando las variables independientes incluyen medidas tanto continuas como categóricas y la variable de respuesta es categórica como por ejemplo (verdadero/falso). Este modelo estima la probabilidad de que los resultados se presenten en una función exponencial de variables independientes”. Es así como se concluye nuestro estudio, mostrando los sucesos posibles entre dos variables dependientes:

Formulas:

$$\Pr(y = 1|x) = \frac{\exp(b_0 + \sum_{i=1}^n b_i x_i)}{1 + \exp(b_0 + \sum_{i=1}^n b_i x_i)}$$

$$\frac{\Pr(y = 1|x)}{1 - \Pr(y = 1|x)} = \exp\left(b_0 + \sum_{i=1}^n b_i x_i\right)$$

$$\log\left(\frac{\Pr(y = 1|x)}{1 - \Pr(y = 1|x)}\right) = b_0 + \sum_{i=1}^n b_i x_i$$

$$\log\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_i X_i$$

$$y = a + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

$$y = \Pr(x) + \varepsilon \begin{cases} y = 1 \Rightarrow \varepsilon = 1 - \Pr(x) \\ y = 0 \Rightarrow \varepsilon = -\Pr(x) \end{cases}$$

$$\frac{\exp(\beta_{i,0} + \beta_1 X_1 + \beta_{i,2} X_2 + \dots + \beta_{i,k} X_k)}{1 + \exp(\beta_{i,0} + \beta_1 X_1 + \beta_{i,2} X_2 + \dots + \beta_{i,k} X_k)}$$

Para este análisis elegimos como variable dependiente tenemos:

*Y = “método copiar y pegar es el adecuado al presentar un trabajo de investigación académica”*

Como variables independientes:

*X<sub>1</sub> = El uso de fuentes bibliográficas que utiliza para investigar*

*X<sub>2</sub> = Motivos por los que realiza un trabajo de investigación*

*X<sub>3</sub> = Horas dedicadas a realizar un trabajo de investigación*

*X<sub>4</sub> = Tipo de fuentes bibliográficas confiables para hacer un trabajo de investigación*

*X<sub>5</sub> = Tipo de bases científicas para hacer sus trabajos de investigación.*

*X<sub>6</sub> = Utilización del tiempo necesario para leer y analizar la información que se busca.*

*X<sub>7</sub> = Adquirir conocimiento adecuado al finalizar el trabajo de investigación*

*X<sub>8</sub> = Utilizar citas bibliográficas en su trabajo de investigación académica.*

El método que se utiliza en regresión logística, es el *Método Atrás de WALD*, que consiste en seleccionar automáticamente por pasos hacia atrás. Se selecciona hacia atrás porque se desea que el modelo incluya en un principio todas las variables independientes y vaya quitando variables en cada paso hasta solo quedar las variables explicativas, esto es posible ya que la variables que se eliminan son aquellas que no cumplen estadísticamente con el grado de significancia ( $p < 0.05$ ).

Tabla 5: Variables que se van eliminando durante los pasos hacia atrás de WALD

	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)		B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Inferior	Superior							Inferior	Superior
<b>Paso 1*</b>	FuentesUtilizar		1,425	3	,700											
	FuentesUtilizar(1)	,110	,706	,024	1	,876	1,117	,280	4,458							
	FuentesUtilizar(2)	,376	,392	,144	1	,705	1,457	,208	10,186							
	FuentesUtilizar(3)	-,232	,683	,115	1	,734	,793	,208	3,023							
	Motivos_hacer_tr		1,044	3	,790											
	Motivos_hacer_tr	-,351	,372	,889	1	,346	,704	,339	1,460							
	Motivos_hacer_tr	-,358	,464	,595	1	,441	,639	,282	1,736							
	Motivos_hacer_tr	-,269	,558	,233	1	,630	,764	,256	2,279							
	FuentesdeInformac		1,386	4	,847											
	FuentesdeInformac	,324	,739	,164	1	,686	1,382	,288	6,622							
	FuentesdeInformac	-,146	,459	,101	1	,751	,864	,352	2,124							
	FuentesdeInformac	,267	,453	,348	1	,555	1,306	,538	3,171							
	FuentesdeInformac	,166	,611	,074	1	,786	1,180	,357	3,906							
	BasesdeConsultas		2,373	4	,667											
	BasesdeConsultas	-,128	,372	,017	1	,895	,880	,131	5,908							
	BasesdeConsultas	,368	,334	1,075	1	,300	2,633	,422	16,424							
	BasesdeConsultas	-,158	,829	,036	1	,849	,854	,168	4,335							
	BasesdeConsultas	-,1398	1,260	1,231	1	,267	,247	,021	2,919							
	LeeryAnalizar(1)	1,526	,812	3,535	1	,060	4,598	,937	22,561							
	UsodeCitas(1)	,175	,608	,083	1	,773	1,192	,362	3,922							
	Constante	-,860	1,226	,431	1	,483	,423									
<b>Paso 2*</b>	FuentesUtilizar		1,771	3	,621											
	FuentesUtilizar(1)	,132	,695	,036	1	,849	1,141	,292	4,459							
	FuentesUtilizar(2)	,384	,387	,151	1	,638	1,467	,212	10,161							
	FuentesUtilizar(3)	-,256	,674	,145	1	,704	,774	,207	2,838							
	Motivos_hacer_tr		,953	3	,813											
	Motivos_hacer_tr	-,332	,371	,802	1	,370	,718	,347	1,483							
<b>Paso 3*</b>	FuentesUtil		2,024	3	,567											
	FuentesUtil	,128	,691	,034	1	,853	1,137	,293	4,405							
	FuentesUtil	,402	,375	,170	1	,680	1,435	,221	10,036							
	FuentesUtil	-,271	,667	,164	1	,685	,763	,206	2,822							
	BasesdeCon		2,153	4	,708											
	BasesdeCon	-,046	,340	,002	1	,361	,355	,151	6,027							
	BasesdeCon	,328	,875	1,123	1	,289	2,529	,455	14,067							
	BasesdeCon	-,049	,803	,004	1	,351	,352	,197	4,538							
	BasesdeCon	-,1254	1,242	1,019	1	,313	,285	,025	3,255							
	LeeryAnaliza	1,505	,762	3,902	1	,048	4,504	1,012	20,052							
	UsodeCitas(	,230	,596	,148	1	,700	1,258	,391	4,046							
	Constante	-,1042	1,122	,863	1	,353	,353									
<b>Paso 4*</b>	FuentesUtil		2,035	3	,565											
	FuentesUtil	,042	,692	,004	1	,951	1,043	,269	4,048							
	FuentesUtil	,223	,342	,056	1	,813	1,250	,197	7,928							
	FuentesUtil	-,358	,669	,287	1	,592	,639	,188	2,531							
	LeeryAnaliza	,875	,467	3,517	1	,061	2,399	,361	5,985							
	UsodeCitas(	,198	,587	,114	1	,736	1,219	,385	3,854							
	Constante	-,323	,897	,130	1	,719	,724									
<b>Paso 5*</b>	FuentesUtil		2,069	3	,558											
	FuentesUtil	,052	,689	,006	1	,340	1,053	,273	4,068							
	FuentesUtil	,226	,340	,058	1	,810	1,253	,199	7,908							
	FuentesUtil	-,354	,667	,282	1	,595	,702	,190	2,531							
	LeeryAnaliza	,884	,466	3,539	1	,058	2,420	,371	6,031							
	Constante	-,152	,737	,043	1	,836	,859									
<b>Paso 6*</b>	LeeryAnaliza	,877	,427	4,211	1	,040	2,403	1,040	5,550							
	Constante	-,310	,397	,610	1	,435	,733									

La tabla 9, nos muestra que por el Método de Wald hacia atrás obtuvimos 6 pasos de las cuales en el primer paso ingresan todas las variables y conforme vaya avanzando se van eliminando las variables que tienen un error estándar mayor que 1 o aquellas que tienen un ODDS RATIO (OR) muy grande o cercano a cero, es así que las variables que se eliminaban por cada paso son:

Variables eliminadas en el paso 2: Fuentes de Información

Variables eliminadas en el paso 3: Motivos hacer trabajos

Variables eliminadas en el paso 4: Bases de Consultas

VARIABLES ELIMINADAS EN EL PASO 5: USO DE CITAS

VARIABLES ELIMINADAS EN EL PASO 6: FUENTES A UTILIZAR

Finalmente en el paso 6 quedan las variables que serán parte de nuestro modelo, en este caso solo quedo una variable (Leer y Analizar), por lo que el modelo ajustado quedaría de la siguiente forma.

Tabla 6: Variables que forman parte de la ecuación

	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Inferior	Superior
Leer y Analizar(1)	,877	,427	4,211	1	,040	2,403	1,040	5,550
Constante	-,310	,397	,610	1	,435	,733		

De acuerdo a los resultados de la tabla 10, la probabilidad de cometer plagio es:

Formulas:

$$\Pr(Y = 1|x) = \frac{\exp(b_0 + \sum_{i=1}^n b_i x_i)}{1 + \exp(b_0 + \sum_{i=1}^n b_i x_i)}$$

$$\Pr(\text{Plagio} = 1) = \frac{1}{1 + \exp(0.310 - 0.877 X_6)}$$

$$\Pr(\text{Plagio} = 1) = \frac{1}{1 + \exp(0.310 - 0.877 (\text{Leer y Analizar}))}$$

Lo que se puede interpretar del resultado de la fórmula es que, la probabilidad de cometer plagio académico, sabiendo que el estudiante no lee ni analiza la información que es parte de su investigación académica es de 0.638 (63.8%).

En lo que queda abierto para futuras investigaciones se quiere hacer mención a la información de que las autoridades de las Universidades en Ecuador, asuman el reto de plantear soluciones a la situación del plagio académico, fomentar y concientizar en los estudiantes valores de honestidad, responsabilidad, respeto y equidad.

## CONCLUSIONES

A partir de los resultados obtenidos se destaca que el uso de internet y la falta de análisis y lectura de los textos que revisa el estudiante, se ha convertido en la fuente principal para que los alumnos cometan plagio en sus trabajos académicos.

La falta de conocimiento de fuentes de información confiable tales como revistas electrónicas, artículos científicos, bases de datos indexadas, hace que el alumno no comprenda la importancia de hacer un trabajo honesto.

El uso frecuente de sitios web no confiables como el rincón del vago, buenas tareas, entre otros hace que el alumno se apodere de información que no es investigada previamente, lo que hace que sus trabajos de investigación académica tengan la calidad de un buen trabajo.

#### BIBLIOGRAFÍA

- Barrón, A. (2010). Detección automática de plagio, una breve reseña. Grupo Fivas.
- Bisang, R. (1995). Las actividades de investigación en las Universidades Nacionales. Buenos Aires: SPU.
- Buitrago J. (2004). "Fraude y engaño en la investigación biomédica". Colombia Med; 35: 93-100.
- Hativa, N. (2000), Teaching for Effective Learning in Higher Education, Dordrecht/ Boston/ London, Kluwer Academic Publishers.
- Hernández, R. (2008). Metodología de la investigación. México: McGrawHill cuarta edición, p. 560.
- Huamani, C., Dulanto-Pizzorni, A., & Rojas-Revoredo, V. (2008). `Copiar y pegar` en investigaciones en el pregrado: haciendo mal uso del Internet. Anales de la Facultad de Medicina. Recuperado 6 de agosto de 2014, a partir de <http://www.redalyc.org/resumen.oa?id=37911344010>
- Khan, M. G. M.; Najmussehar and Ahsan, M .J. (2005): <<Optimum Stratification for Exponential study Variable under Neyman Allocation>>. Journal of the Indian Society of Agricultural Statistics 59 (2): 146-150.
- Kibler, W.L. (1993). Academic dishonesty: A student development dilemma. NASPA Journal, 30, 252-267
- Londoño FJL. (1995). Metodología de la investigación epidemiológica. Ed. Universidad de Antioquia, Medellín.
- Molinero, Luis M. La regresión logística (I). Bioestadística. [consultado 2/2008]. Disponible en: <http://www.seh-lelha.org/rlogis1.htm>: Febrero.
- Pértegas Díaz S, Pita Fernández, S. Determinación del tamaño muestral para calcular la significación del coeficiente de correlación lineal de Pearson. Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario Juan Canalejo. A Coruña (España): cad aten primaria 2001; 2002; 9: 209-211.
- Sánchez, J. (2000). Nuevas tecnologías de la información y la comunicación para la construcción del aprender. Santiago de Chile: Centro Zonal Universidad de Chile, Proyecto Enlaces.
- Sánchez, M.P. (2006). Nociones básicas en materia de propiedad intelectual. Periférica.

