
La aplicación de tablas de contingencia para relacionar variables como factores maternos

The application of contingency tables to relate variables as maternal factors



Carrillo, Ronnie E.; Arias, José E.; Mejillones, Erick L.; Naranjo, Pedro A.

Ronnie E. Carrillo

ronnie.carrillod@gmail.com

Universidad de Guayaquil, Ecuador., Ecuador

José E. Arias

jose.arias@gmail.com

Universidad de Guayaquil, Ecuador., Ecuador

Erick L. Mejillones

erick.limonm@gmail.com

Universidad de Guayaquil, Ecuador., Ecuador

Pedro A. Naranjo

pedro.naranjoi@gmail.com

Universidad de Guayaquil, Ecuador., Ecuador

Investigación, Tecnología e Innovación

Universidad de Guayaquil, Ecuador

ISSN: 1390-5147

ISSN-e: 2661-6548

Periodicidad: Anual

vol. 12, núm. 12, 2020

revistaiti@ug.edu.ec

Recepción: 03 Diciembre 2019

Aprobación: 21 Octubre 2020

URL: <http://portal.amelica.org/ameli/jatsRepo/593/5932566006/index.html>

Resumen: El objetivo de este trabajo es identificar el riesgo de defunción fetal relacionado a los factores maternos de las mujeres que habitan en las zonas urbanas y rurales. El problema radica en la poca importancia que toman las mujeres en las distintas zonas (ya sea urbana o rural) de estos factores, al momento de determinar la causa de defunción del organismo, debido a que estas no lo consideran como algo influyente, lo cual tiene como principal consecuencia una escasa información sobre si alguno de estos acontecimientos fue el responsable o no del fallecimiento del organismo. Para poder resolver este problema se emplean ciertos algoritmos que nos ayudan a correlacionar dichos factores aplicando tablas de contingencia y chi cuadrado, adicionalmente se representaran en el análisis causa raíz para facilitar el entendimiento de dichos factores. se determinó que uno de los principales factores que afectan al feto en zonas urbanas esto puede ser debido la contaminación ambiental en áreas con mucho smoke, lo cual causa la hipoxia del feto (4,89) a diferencia de las zonas rurales donde el principal factor es el déficit de conocimiento acerca del embarazo y la escasez económica por otro lado es contradictorio que las zonas urbanas posean más control prenatal y asistencia médica no obstante es la zona con más muertes presentada con 1558 casos registrados.

Palabras clave: chi-cuadrado, área residencial, mortalidad fetal, defunción.

Abstract: The objective of this work is to identify the risk of fetal death related to maternal factors of women living in urban and rural areas. The problem lies in the little importance that women take in the different areas (whether urban or rural) of these factors, when determining the cause of death of the organism, because they do not consider it as something influential, which the main consequence is little information on whether or not any of these events was responsible for the death of the organism. In order to solve this problem, certain algorithms are used that help us to correlate these factors by applying contingency tables and chi square, additionally they will be represented in the root cause analysis to facilitate the understanding of these factors. It was determined that one of the main factors that affect the fetus in urban areas, this may be due to environmental pollution in areas with a lot of smoke, which causes hypoxia of the fetus (4.89), unlike in rural areas where the main factor It is the lack of knowledge about pregnancy and the economic scarcity, on

the other hand, it is contradictory that urban areas have more prenatal control and medical assistance, however, it is the area with the most deaths with 1558 registered cases.

Keywords: chi-square, residential area, fetal mortality, death.

INTRODUCCIÓN

La importancia de ser madre es un privilegio que tienen las mujeres en base al instinto materno y el amor que genera expectativa e ilusiones que van fomentando día a día la vida de la futura madre y su hijo, sin embargo, muchas mujeres no llegan a cumplir sus sueños debido a factores que involucran la muerte prematura de su hijo. El problema persiste en el análisis de estos factores y si se consideran influyentes como causa de la defunción del feto, no es posible analizarlos debido a las grandes muestras que se tienen de las diferentes zonas rurales y urbanas del Ecuador en consecuencia resulta muy difícil analizar cada uno de los factores para determinar si alguno fue influyente o no.

Trabajos Relacionados

Incluso después del parto pueden presentarse ciertas adversidades que pueden terminar con la vida del individuo tal como se evidencia en (Hernández J., Hernández M. G., Herrera M., Ló

Incluso después del parto pueden presentarse ciertas adversidades que pueden terminar con la vida del individuo tal como se evidencia en (Hernández J., Hernández M. G., Herrera M., López M., Rodríguez J. G. y Guerrero A., 2013), que desarrolló un estudio descriptivo sobre recién nacidos que presentaban dificultad respiratoria, con el fin de conocer la frecuencia de este síndrome y establecer los factores de riesgos asociados, se implementó una correlación de Spearman para poder observar el grado de asociación entre las variables expulsivo prolongado, ruptura prematura de membrana y traumatismo durante el parto. El estudio obtiene como resultado la frecuencia de bronco aspiración en relación a la edad gestacional de los recién nacidos. A pesar de ello este trabajo si logra probar la relación entre las variables, no obstante, el trabajo realizado en este paper utiliza un análisis de tablas de contingencia la cual permite el manejo de grandes cantidades de información incluso de variables cuantitativas y cualitativas cosa que el método de Spearman no es capaz de realizar.

La defunción fetal puede ocurrir en cualquier momento y en un alto porcentaje se desconocen sus causas tal como se refleja en (Tinedo M. J., Santander P. F., Alonso F. J., Herrera H. A., S. C. y Díaz M., 2016) donde realizaron un estudio retrospectivo sobre el diagnóstico de muerte fetal los datos fueron obtenidos a través de la revisión de historias clínicas. Esta investigación se realiza mediante un análisis de correlación de Pearson, el estudio obtiene como resultado la tasa de mortalidad con respecto a cada nacimiento y abortos dentro del hospital y la relación a la zona residencial sin embargo en el trabajo presentado por Herrera H y Adrián; Si demuestra que existe una relación entre las dos variables, pero este paper se diferencia porque se aplica el estadístico de chi cuadrado para poder demostrar si existe una relación entre las variables mediante el descarte de hipótesis.

La muerte está presente en todas las edades sin excepción, así como se evidencia en (González Galbán H. y Herrera León L. I., 2015) donde aplicaron un análisis hacia algunos factores considerados influyentes en la mortalidad de los infantes, para poder determinarlo se aplicó el método demográfico de la tabla de vida. El estudio obtiene como resultado la mortalidad de los infantes con respecto al orden del nacido vivo. A pesar de que en el trabajo de González y Herrera muestra logran representar la tasa de muerte infantil, a diferencia de este paper donde además de mostrar el índice de muerte se evidencian los factores influyentes mediante la

representación simultanea de dos variables observadas en la misma población y además se realizó una prueba de hipótesis que compara la distribución observada de los datos.

Trabajar con grandes cantidades de información es muy complicado más si no se usa ningún tipo de aplicación informática tal como plantea (Delgado M., Muñoz A., Orejuela L. y Sierra C. H., 2003), El cual realizo un análisis estadístico de los datos mediante una base de datos en Microsoft Excel para posteriormente ser trasladados a SPSS una vez ahí se aplicó la prueba de Chi cuadrado para determinar la asociación entre cada variable y su riesgo de mortalidad , el estudio tiene como resultado la asociación entre la historia materna y el riesgo de mortalidad con el fin de poder establecerlos como factores de riesgo , pese a que SPSS si demuestra la relación en el trabajo propuesto por Delgado, Mario y Muñoz, Alejandro , en este paper hace uso del programa Rstudio debido a que posee una gran variedad de funcionalidades para el tratamiento y el análisis de datos además porque es un software de acceso libre y es capaz de trabajar con grandes cantidades de datos estadísticos gracias a su versatilidad.

MATERIALES Y MÉTODOS

Cuando se posean dos variables siempre surgirá la duda de si están relacionadas o no y si lo llegasen a estar que tan relacionadas están, para conocer y comprobar esto en una tabla de contingencia aplicamos el estadístico de chi-cuadrado el cual nos ayudara a despejar esta duda, además sabemos lo duro que es ser madre y más lo duro que tiene que ser perder a un hijo por eso, mediante un diagrama de Ishikawa determinaremos estas causas para lograr que esta trágica consecuencia descienda lo más que se pueda.

Las tablas de contingencia se integran por filas o renglones y columnas, las cuales forman celdas o casillas en donde se registra la frecuencia absoluta o relativa de cada una de las categorías analizadas. En la tabla de contingencia, entonces, encontramos en una celda la información de la frecuencia que existe en los casos en que se comparte la unidad de análisis (Sanchez Ramoz M. A., 2005).

Por otra parte (Rouquette J. Ó. y Saleme M., 2000.), define como: "Una tabla que clasifica datos de acuerdo con dos o más categorías, relacionados con cada una de dos variables cualitativas, que pueden ser o no estadísticamente independientes.

TABLA 1.
Tabla de contingencia de forma general entre dos variables ya sean estas cualitativa y cualitativa o cualitativa y cuantitativa.

<i>X/Y</i>	<i>Y₁</i>	<i>Y₂</i>	
<i>X₁</i>	<i>n₁₁</i>	<i>n₁₂</i>	<i>n₁₁ + n₁₂</i>
<i>X₂</i>	<i>n₂₁</i>	<i>n₂₂</i>	<i>n₂₁ + n₂₂</i>
Total	<i>n₁₁ + n₂₁</i>	<i>n₁₂ + n₂₂</i>	<i>N</i>

X₁ & Y₁ numero de individuos para el cual toma valor x & y variables

$$(1)$$

$\sum_{k=1}^s n_{11} + n_{21}$: es la sumatoria de la columna para el valor de *Y₁* (2)

$$\sum_{k=1}^s n_{11} + n_{21} : \text{es la sumatoria de la columna para el valor de } Y_1 \text{ (3)}$$

Estadístico chi-cuadrado

La prueba del χ^2 se usa para variables de distintos niveles de medición, incluyendo las de menor nivel, que son las nominales. Sirve para determinar si los datos obtenidos de una sola muestra presentan variaciones estadísticamente significativas respecto de la hipótesis nula (Nievas F., 2011).

También debemos tener en cuenta lo que nos aclara (De la Fuente F. Santiago, 2016), que dice Es necesario que las frecuencias esperadas de las distintas modalidades no sean inferiores a cinco. Si alguna modalidad tiene una frecuencia esperada menor que cinco se agrupan dos o más modalidades contiguas en una sola hasta conseguir que la frecuencia esperada sea mayor que cinco.

Para analizar la relación de dependencia y/o independencia entre dos variables cualitativas, se utiliza la prueba Chi Cuadrado de Independencia. Esta prueba permite determinar si existe una relación entre dos variables categóricas (cualitativas). Es necesario resaltar que esta prueba indica si existe o no una relación entre las variables, pero no señala el grado o el tipo de relación; es decir, no indica el porcentaje de influencia de una variable sobre la otra o la variable que causa la influencia (Tinoco Gómez O., 2008).

$$\chi^2 = \sum_{ij} \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

n_{ij} =frecuencia observada para la ij -ésima casilla.

e_{ij} =frecuencia esperada para la ij -ésima casilla.

Hipótesis nula y de trabajo

Al comparar variables se deben de formular dos hipótesis entre ella la de trabajo que será donde supongamos que existe una relación, su contra será la hipótesis nula tal y como nos dice (Nievas F., 2011), De acuerdo a la hipótesis nula (H_0) las variaciones en la variable independiente no tienen correspondencia con las variaciones que pudiere haber de la variable dependiente. Es decir que existe “independencia estadística”. Las variaciones que pudiese encontrarse se deberían a factores aleatorios, ajenos a la variable independiente.

También nos dice Inzunza (INZUNSA CAZARES S. y JIMÉNEZ RAMÍREZ J. VIDAL, 2013), prueba de hipótesis es una regla de comportamiento inductivo que permite elegir entre una hipótesis nula y una hipótesis alternativa. La evidencia de los datos obtenidos puede conducir a no rechazar la hipótesis nula, lo cual no implica que ésta sea cierta se considera que la principal diferencia entre estas dos teorías no radica en los cálculos, sino en las concepciones y el razonamiento subyacente.

H_0 : Las variables en filas y columnas no están asociadas.

H_1 : Las variables en filas y columnas están asociadas.

$$H_0 \text{ cuando } \chi^2 \text{ experimental} > \chi^2 \text{ critico} \tag{5}$$

Grados de libertad

En el estudio de De La Cruz (De La Cruz Oré J. L., 2013), nos dice que la suma de los valores de las desviaciones de los valores individuales con respecto a su media es igual a cero, hecho que puede demostrarse. Si se conocen los $n-1$ valores de los valores a partir de la media, entonces se conoce el n -ésimo valor, ya que queda determinado automáticamente debido a la restricción de 3 que todos los valores de n sumen cero.

En la investigación de (Universidad de Granada, 2013), nos dice que el grado de libertad es un estadístico calculado sobre un conjunto de datos se refiere al número de cantidades independientes que se necesitan en su cálculo, menos el número de restricciones que ligan a las observaciones y el estadístico. El número de grados de libertad del estadístico Chi-cuadrado se calcula de la siguiente forma:

- Se calcula, en primer lugar el número de sumandos, es decir $m \times n$, siendo n y m el número de filas y número de columnas en la tabla.
- A esta cantidad se debe restar el número de restricciones impuestas a las frecuencias observadas. Observamos que podemos cambiar todas las frecuencias de la tabla sin cambiar los totales por filas y columnas, excepto los datos en la última fila y columna de la tabla, pues una vez que fijemos todos los valores excepto estos, quedan automáticamente fijados. Por tanto, si la tabla tiene m filas y n columnas, el número de grados de libertad es

$$(k - 1) * (m - 1) \text{ o } (filas - 1) * (col - 1) \tag{6}$$

Cálculo de las frecuencias esperadas

Para el cálculo de las frecuencias esperadas (f_e) (Nievas F., 2011), nos vuelve a colaborar diciendo (f_e) que, vienen dadas por la hipótesis nula (H_0), pero no siempre se puede establecer de manera inmediata. Esto solo es posible cuando trabajamos con una variable, pero cuando tenemos cuadros de doble entrada la forma de establecer el valor de la frecuencia esperada de cada celda es el siguiente:

TABLA 2.
Tabla de contingencia de la cual saldrá la tabla de frecuencias esperadas.

<i>X/Y</i>	<i>Y₁</i>	<i>Y₂</i>	<i>Marginal Y1/Y2</i>
<i>Categoría A</i>	<i>a</i>	<i>b</i>	<i>(a + b)</i>
<i>Categoría B</i>	<i>d</i>	<i>e</i>	<i>(d + e)</i>
<i>Categoría C</i>	<i>g</i>	<i>h</i>	<i>(g + h)</i>
<i>Marginal A/B/C</i>	<i>(a + d + g)</i>	<i>(b + e + h)</i>	<i>N</i>

Se debe aplicar la siguiente fórmula para obtener los valores esperados.

$$e_{i,j} = \frac{f_i \times f_j}{n} \tag{7}$$

Para la celda a

$$\frac{(a + d + g)(a + b)}{N} \tag{8}$$

Para la celda b

$$\frac{(b + e + h)(a + b)}{N} \tag{9}$$

Y así con todas las celdas que se posea en la tabla y se obtendrá la Tabla 3.

TABLA 3.
Tabla general de frecuencias esperadas obtenida una vez ya aplicada la formula a cada celda.

<i>X / Y</i>	<i>Y₁</i>	<i>Y₂</i>	<i>Marginal Y1/Y2</i>
<i>Categoría A</i>	<i>F.E a</i>	<i>F.E b</i>	<i>(a + b)</i>
<i>Categoría B</i>	<i>F.E d</i>	<i>F.E e</i>	<i>(d + e)</i>
<i>Categoría C</i>	<i>F.E g</i>	<i>F.E h</i>	<i>(g + h)</i>
<i>Marginal A/B/C</i>	<i>(a + d + g)</i>	<i>(b + e + h)</i>	<i>N</i>

Estadístico Teórico

Para obtener el valor teórico primero debemos tener en cuenta el grado de libertad y el nivel de significación α pero este es arbitrario y se fija de antemano (usualmente entre 0.01 y 0.10, siendo el más usado el de 0.05).luego en la tabla de chi-cuadrado se procede a elegir el número de grado de libertad y el nivel de significancia de 0.05 y el valor será el número que se encuentre en la intersección de ambas (D. P., Cortés Bernal y Cardona Rivas D., 2014).

$$X^2_{\alpha (k-1) (m-1)} \tag{10}$$

TABLA 4.
Parte seleccionada de la tabla de chi cuadrado con los valores de significancia hasta 0.05.

V/P	0.001	0.0025	0.005	0.01	0.025	0.05
1	10.8274	9.1404	7.8794	6.6349	5.0239	3.8415
2	13.8150	11.9827	10.5965	9.2104	7.3778	5.9915
3	16.2660	14.3202	12.8381	11.3449	9.3484	7.8147
4	18.4662	16.4238	14.8602	13.2767	11.1433	9.4877
5	20.5147	18.3854	16.7496	15.0863	12.8325	11.0705

Contraste de independencia de chi-cuadrado

Una vez obtenida la tabla de frecuencias esperadas se procede a aplicar el estadígrafo de contraste que se presenta a continuación a cada uno de las celdas con dicha frecuencia

$$x^2 = \sum_{ij} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \quad (11)$$

n_{ij} =frecuencia observada para la ij-esima casilla.

e_{ij} =frecuencia esperada para la ij-esima casilla.

Se elige un estadístico de contraste este será el que determine cual hipótesis se debe de escoger si la nula (H_0) o la de trabajo (H_1). Pero primero hay que calcular el valor experimental como lo presenta (De la Fuente F. Santiago, 2016) que será la sumatoria de los valores calculados con la fórmula 5.

$$X_{exp}^2 = \sum_{i=1}^k \sum_{j=1}^m \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \quad (12)$$

Una vez que calculado el valor experimental y el valor del estadístico teórico:

Se aceptará H_0 si:

$$X_{exp}^2 = \sum_{i=1}^k \sum_{j=1}^m \frac{(n_{ij} - e_{ij})^2}{e_{ij}} < x_{\alpha(k-1),(m-1)}^2 \quad (13)$$

Se rechaza H_0 si:

$$X_{exp}^2 = \sum_{i=1}^k \sum_{j=1}^m \frac{(n_{ij} - e_{ij})^2}{e_{ij}} > x_{\alpha(k-1),(m-1)}^2 \quad (14)$$

Por ende, si se cumple la fórmula (7), las variables serán independientes caso contrario será dependiente una de la otra.

Diagrama de Ishikawa o causa raíz

Esta herramienta no ofrece respuesta a una pregunta y en el momento de generar el diagrama causa-efecto, normalmente se ignora si estas causas son o no responsables de los efectos. Por otra parte, un diagrama causa-efecto bien organizado sirve como vehículo para ayudar a los equipos a tener una concepción común de un problema complejo, con todos sus elementos y relaciones claramente visibles a cualquier nivel de detalle requerido (Zapata Carlos M. y Villegas Sandra M., 2006).

Por otro lado, Gandara (Gándara González F. d. J., 2014), el Diagrama de Causa y Efecto por su forma recibe el nombre de “esqueleto de pescado”, en el que la espina dorsal es el camino que conduce a la cabeza del pescado que es donde se coloca el problema que se desea analizar; las espinas o flechas que la rodean indican las causas y sub-causas que lo provocan como se muestra en la figura.

Diagrama de Ishikawa o de Causa Raíz entorno a la muerte Fetal



FIGURA 1.

Diagrama de Ishikawa o de causa raíz. Se observa los factores vinculados después del análisis del estudio previo que inciden en la muerte fetal.

Análisis del grafico

Las principales causas que afectan y generan perturbación a la madre en el proceso de embarazo se encuentra en el poco conocimiento acerca de los riesgos y cuidados que se debe tener durante este periodo además del entorno en el que se vive afecta mucho al proceso de cuidado, no obstante, si bien el problema es más grave en las zonas urbanas en la zona rural hace falta la concientización acerca de los controles prenatales, así como un buen seguimiento de un médico especializado.

Solución

- Se propone campañas de concientización en las zonas rurales sobre la importancia del seguimiento médico, así también como de información acerca de los cuidados del proceso materno no solo para la madre si no para los familiares debido a que estos la rodean.
- También mejor acceso a la salud publica en estas zonas para generar un mejor control y evitar más muertes.

Antecedentes

Según el trabajo hecho por (Zoila Valladares B., Vanessa García D., Vanesa Buján C., Couceiro Emilio N. y Nicolás López Carlos R., 2013) anualmente en el mundo se producen cerca de 4 millones de muertes fetales, y otro par de millones no son contados, una de las causas principales en algunos países es su falta de asistencia sanitaria, y generalmente estas defunciones ocurren en países en pleno proceso de desarrollo, mientras que en países desarrollados esta tasa disminuye considerablemente. El estudio fue descriptivo retrospectivo se analizaron todos los casos de muerte fetal tardía en gestaciones únicas acontecidas en el Complejo Hospitalario Universitario de Vigo (2005- 2010). Y como resultado se obtuvo 56 casos de muerte

fetal tardía. A pesar de los costes derivados de dicho estudio, la información obtenida puede ayudar a estimar el riesgo de recurrencia y proporcionar orientación para embarazos futuros.

Por otra parte, Gonzales (González Galbán H. y Herrera León L. I., 2015), nos dice que la incidencia de la mortalidad cercana al nacimiento está asociada potencialmente a las características biológicas y demográficas de la madre durante su historia reproductiva. A fin de analizar los señalados aspectos, se presenta el actual artículo que tiene como objetivo general el valorar si algunos de los factores biodemográficos que aquí son considerados constituyen condicionantes para la ocurrencia de la mortalidad infantil y/o para el resto de los infantes menores de 5 años.

También Cevallos (A. G. R. J. M. A. C. G. S. D. V. M. K. B. R. R. M. A. Cevallos Torres F. Lorenzo, 2018) en su investigación busco evaluar el efecto de la eritropoyetina en la corrección de la anemia en pacientes con insuficiencia renal crónica que reciben tratamiento de Hemodiálisis, los datos recogidos corresponden a una base de datos perteneciente a la unidad de Hemodiálisis CENAG SA de la ciudad de Guayaquil, durante el periodo comprendido entre junio 2010 a junio 2014. Para este estudio se ha realizado un análisis estadístico de regresión lineal múltiple y análisis de varianza, con la finalidad de evaluar mejoras en los niveles de hemoglobina en respuesta al uso de la eritropoyetina.

En común con nuestro trabajo, está el trabajo de (A. G. R. L. L. D. R. Cevallos Torres Lorenzo, 2017), el cual busca los factores de incidencia en un tema específico solo que este trabajo tiene 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.

Programa

Rstudio es un conjunto de herramientas integradas diseñadas para ayudarlo a ser más productivo con R. Incluye una consola, un editor de resaltado de sintaxis que admite la ejecución directa de código y una variedad de herramientas robustas para trazar, ver el historial, depurar y administrar su espacio de trabajo (Rstudio).

Algoritmos

```

Val_max=11
while(x<11){

  edad_16_19URBANA=subset(DATOS_SIN_99,edad_mad>=v1&edad_mad<v1+3 & area_res=="Urbana")

  urbana=length(edad_16_19_URBANA$edad_mad)
  tt=length(edad_16_19$edad_mad)
  if(bandera>0){
    vector_RURAL=c( vector_RURAL,rural)
    vector_Urbana=c(vector_Urbana,urbana)
    vector_TT=c(vector_TT,tt)
  }
  else
  {
    vector_RURAL=rural
    vector_Urbana=urbana
    vector_TT=tt
  }
  x=x+1
}

Tabla_Edad=data.frame(Urbana=vector_Urbana,Rural=vector_RURAL)
    
```

ALGORITMO 1.

El siguiente bloque nos permite generar la tabla de contingencia entre las variables que se desee.

```

totalY=c(1418,278, 5, 117)
totalX=c(1558,260,1818)
nx=1
ny=3

while (nx<=4) {
  aux=(totalY[nx]*totalX[2])/totalX[3]
  if(nx==1) {

    urbana=c(aux)

  }
  else{
    urbana=c(urbana,aux)
  }
  nx=nx+1
}
    
```

ALGORITMO 2.

Algoritmo en R studio para obtener los valores de la frecuencia esperada de la tabla de frecuencias esperadas.

```
chisq.test(tabla_area)
```

ALGORITMO 3.

Para obtener el valor del chi-cuadrado hacemos uso de la siguiente línea de código Aplicada a la tabla de contingencia ya generada.

```
barplot(tabla_area)
```

ALGORITMO 4.

La siguiente línea de código nos generará la gráfica de la tabla de contingencia.

Análisis de datos

Para el análisis de los datos proporcionados en la Base de datos del INEC se extrajeron 6 variables las cuales nos ayudaran en el proceso de obtención de resultados las cuales son asistido por, edad de la madre, causa de la mortalidad, controles prenatales, provincia de fallecimiento del feto, asistido por y el área residencial siendo esta el objetivo a comparar con las otras variables para as determinar si la zona sea urbana o rural tienen que ver en los deceso del no nacido, para manejar estos datos que son amplios se usara la herramienta informática de Rstudio la cual de forma fácil y no compleja nos permite generar las tablas de contingencia y la aplicación de chi-cuadrado para hallar alguna relación entre las variables propuestas.

Caso de estudio

Se posee una base de datos proporcionados por el INEC del año 2016, en el cual 1818 personas de las cuales todas son madres que presentaron casos registrados de defunción fetal en las cuales claramente perdieron a sus hijos, pero este problema tiene sus causales los cuales fueron determinados por 41 variables, de las cuales nos centraremos en el área residencial la cual esta subdividida en Urbana y Rural siendo esta el motivo de la inquietud del que está pasando en estas zonas (Cevallos Torres L. y Botto Tobar M. , 2019).

RESULTADOS Y DISCUSIÓN

Variable provincia de fallecimiento vs área residencial

TABLA 5.
Tabla de contingencia de las variables provincia de fallecimiento y area residencial.

PROVINCIA/ÁREA	URBANA	RURAL	TOTAL
GUAYAS	725	27	752
PICHINCHA	302	92	394
OTROS	531	141	672
FRE.MARGINAL	1558	260	1818

En la presente tabla se muestran los datos recabados y tabulados de la población registrada con relación a la provincia donde habitan y el área residencial donde viven siendo el Guayas y Pichincha dos de las provincias con más defunciones fetales registradas y en relación con la zona se evidencia que el área urbana presenta 1558 casos de muertes fetales en comparación a la zona rural.

TABLA 6.
Tabla de frecuencias esperadas obtenida una ves ya aplicada la formula a cada celda.

PROVINCIA/ÁREA	URBANA	RURAL	TOTAL
GUAYAS	644,453245	107,546755	752
PICHINCHA	337,652365	56,3476348	394
OTROS	575,894389	96,1056106	672
TOTAL	1558	260	1818

A partir de la tabla anterior una vez aplicado el cálculo correspondiendo tal como la tabla 3 nos muestra, los nuevos valores serán las frecuencias esperadas y su suma de valores marginales será igual al total de nuestra tabla de contingencia original lo cual nos indica que el cálculo fue concretado con éxito.

A continuación, se plantean las hipótesis a contrastar mediante el chi-cuadrado.

H_0 : La variable provincia de fallecimiento es independiente al área residencial.

H_1 : La variable provincia de fallecimiento si depende del área residencial.

Aplicamos las componentes del chi-cuadrado de la formula (11) y se obtiene el primer valor y así con los valores siguientes.

$$x^2 = \sum_{ij} \frac{(725 - 644.45)^2}{644.45} = 10.067$$

Luego se aplicó el estadístico de contraste de la formula (12) donde nos indica que es la suma de los valores de x^2 .

$$X_{exp}^2 = \sum_{i=1}^k \sum_{j=1}^m \frac{(725 - 644.45)^2}{644.45} = 10.067 + \dots = 121.19$$

Calculamos el valor Teórico de la fórmula (6) pero antes resolvemos la formula (1) Para hallar el valor de los grados de libertad.

$$(fil - 1) * (col - 1) = (3 - 1) * (2 - 1) = 2 * 1 = 2$$

FIG2.

Test de chi-cuadrado realizado en Rstudio donde podemos constatar que si nos coinciden con el calulo matematico.

```
Pearson's Chi-squared test
data: tabla_area
X-squared = 121.19, df = 2, p-value < 2.2e-16
```

$$X_{0.05,2}^2 = 5.99$$

La decisión es: $X_{exp}^2 > X_{0.05,2}^2 = 121.19 > 5.99$ se rechaza H_0 mediante el contraste de chi-cuadrado, por ende, la variable de provincia de fallecimiento si depende del área residencial y si se verán influenciados dependiendo de la provincia y el área en donde viva la madre.

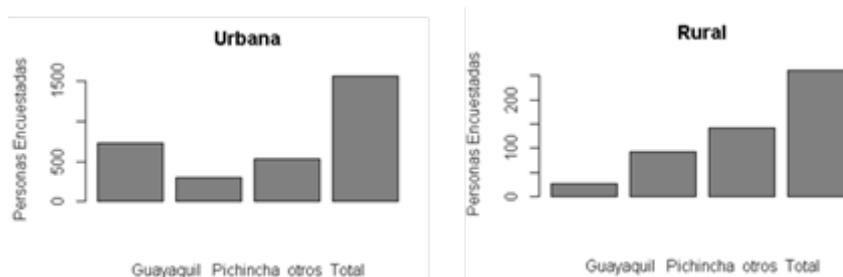


FIGURA 3.

Graficas en Rstudio de la tabla de contingencia con respecto a la provincia y el área residencial.

En la figura 3 se nota que el mayor índice de muertes se presenta en la provincia del Guayas en el área residencial urbana, mientras que en la rural se presencia mayor cantidad en la zona rural la cual se ve que es más notoria en pichincha.

Variable asistido por vs el área residencial

ASISTIDO X /ÁREA	URBANA	RURAL	TOTAL
MEDICO/A	1449	223	1672
OBTETRIZ/OBSTETRA	95	28	123
PARTERO/A NO CALIFICADO/A	2	0	2
OTROS	12	9	21
TOTAL	1558	260	1818

TABLA 7.

Tabla de contingencia de las variables de asistido por y area residencial.

Como se evidencia en la tabla 7 las madres son asistidas con mayor frecuencia por los medicos, tanto en la zona urbana como en la rural, este resultado era de esperace debido a su basto conocimiento, tambien entre la asistencia por un obstetra y de otras areas son asistidos con mas regularidad.

TABLA 8.

Tabla de frecuencias esperadas obtenida unas ves ya aplicada la formula a cada celda.

ASISTIDO X /ÁREA	URBANA	RURAL	TOTAL
MEDICO/A	1432,88	239,119	1672
OBTETRIZ/OBSTETRA	105,409	17,590	123
PARTERO/A NO CALIFICADO/A	1,713	0,286	2
OTROS	17,996	3,0033	21
TOTAL	1558	260	1818

Plantemos nuestras hipótesis que pondremos a prueba, nos permitirá ver la dependencia o independencia de las variables analizadas.

H_0 : La variable asistido por es independiente al área residencial.

H_1 : La variable asistido por si depende del área residencial.

Aplicamos las componentes del chi-cuadrado de la formula (11).

$$x^2 = \sum_{ij} \frac{(1449 - 1432.88)^2}{1432.88} = 0.18$$

Con esta fórmula comparamos la frecuencia obtenida en la tabla de contingencia con la frecuencia que se obtuvo en la tabla de frecuencias esperadas, obteniendo como respuesta 0.18
 Luego se aplicó el estadístico de contraste de la formula (12).

$$X^2_{exp} = \sum_{i=1}^k \sum_{j=1}^m \frac{(1449 - 1432.88)^2}{1432.88} = 0.18 + \dots = 22.761$$

Se procede a calcular el valor experimental o más conocido como estadístico de contraste el cual nos arrojó un valor de 22.761.
 Calculamos el valor Teórico de la fórmula (6) pero antes resolvemos la formula (1)
 Para hallar el valor de los grados de libertad.

$$(fil - 1) * (col - 1) = (4 - 1) * (2 - 1) = 3 * 1 = 3$$

$$X^2_{0.05,3} = 7.814$$

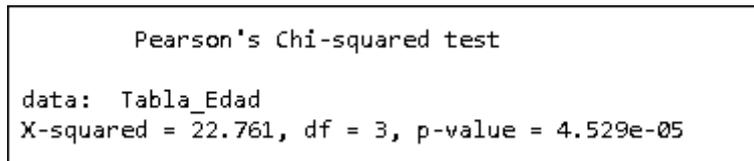


FIGURA 4.

Valor de chi-cuadrado generado a partir de la variable asistido por vs el área residencial.

Como observamos en la figura nos arrojó el mismo resultado que se calculó mediante las formulas anteriores, comprobando así la precisión del software con respecto a nuestras variables analizadas.

La decisión es: $X^2_{exp} > X^2_{0.05,3} = 22.761 > 7.814$ se rechaza H_0 por ende la asistencia que tiene la madre si depende del área residencial urbana y rural.



FIGURA 5.

Graficas en con respecto a la variable asistido por y el área residencial.

Se observa que en la zona urbana y zona rural prefieren en mayoría ser atendidos por un médico que, por especialistas en otras áreas, algo más que notamos es que pocas son las madres que se arriesgan a ser atendidas por algún partero no especializado.

Variable causa de muerte fetal vs el área residencial

TABLA 7.
Tabla de contingencia de las variables de asistido por y area residencial.

ASISTIDO X /ÁREA	URBANA	RURAL	TOTAL
MEDICO/A	1449	223	1672
OBETRIZ/OBSTETRA	95	28	123
PARTERO/A NO	2	0	2
CALIFICADO/A			
OTROS	12	9	21
TOTAL	1558	260	1818

Como se evidencia en la tabla 7 las madres son asistidas con mayor frecuencia por los medicos, tanto en la zona urbana como en la rural, este resultado era de esperace debido a su basto conocimiento, tambien entre la asistencia por un obstetra y de otras areas son asitidos con mas regularidad.

TABLA 8.
Tabla de frecuencias esperadas obtenida unas ves ya aplicada la formula a cada celda.

ASISTIDO X /ÁREA	URBANA	RURAL	TOTAL
MEDICO/A	1432,88	239,119	1672
OBETRIZ/OBSTETRA	105,409	17,590	123
PARTERO/A NO	1,713	0,286	2
CALIFICADO/A			
OTROS	17,996	3,0033	21
TOTAL	1558	260	1818

Plantemos nuestras hipótesis que pondremos a prueba, nos permitirá ver la dependencia o independencia de las variables analizadas.

H₀: La variable asistido por es independiente al área residencial.

H₁: La variable asistido por si depende del área residencial.

Aplicamos las componentes del chi-cuadrado de la formula (11).

$$x^2 = \sum_{ij} \frac{(1217 - 1215.205)^2}{1215.205} = 0.0026$$

Luego se aplicó el estadístico de contraste de la formula (12).

$$X_{exp}^2 = \sum_{i=1}^k \sum_{j=1}^m \frac{(1217 - 1215.205)^2}{1215.205} = 0.0026 + \dots = 3.9542$$

Calculamos el valor Teórico de la fórmula (6) pero antes resolvemos la formula (1)

Para hallar el valor de los grados de libertad en el intervalo de 17-26 se puede observar un valor menor a 5 pero debido a que este está por debajo del 20 % de lo aceptable de chi-cuadrado es válido su valor.

$$(fil - 1) * (col - 1) = (4 - 1) * (2 - 1) = 3 * 1 = 3$$

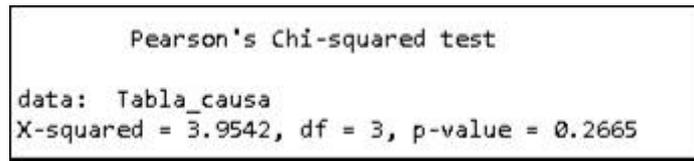


FIGURA 7.

Test de chi-cuadrado de las variables controles prenatales frente al área residencial.

$$X_{0.05,3}^2 = 7.814$$

La decisión es: $\chi_{exp}^2 > \chi_{0.05,3}^2 = 3.9542 < 7.814$ se acepta H_0 como hipótesis nula por ende los controles prenatales no dependen del área residencial urbana y rural.

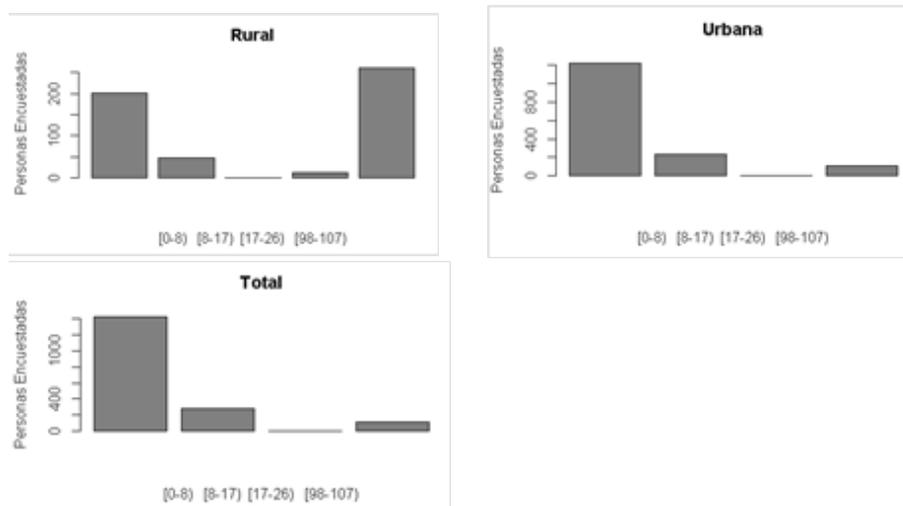


FIGURA 8.

Graficas realizada en Rstudio, muestra la tabla de contingencia entre los controles prenatales y el área residencial.

En la presente figura se puede observar la frecuencia de controles prenatales en relación con las zonas residenciales teniendo mayor frecuencia las zonas urbanas en comparación a las zonas rurales

Variable controles prenatales vs el área residencial

TABLA 13.
Tabla de contingencia de las variables edad de la madre y área residencial.

EDAD/AREA	URBANA	RURAL	TOTAL
13-16	46	3	49
16-19	167	30	197
19-22	241	42	283
22-25	213	34	247
25-28	208	38	246
28-31	191	34	225
31-34	165	19	184
34-37	147	23	170
37-40	89	21	110
40-43	58	10	68
43-46	29	6	35
46-49	3	0	3
TOTAL	1557	260	1817

En la presente tabla se evidencian las edades en las cuales las madres perdieron a sus hijos, podemos observar que la mayor tasa de perdidas está entre mujeres de 19 a 22 años 241 pertenecientes a las zonas urbanas. Podemos deducir gracias al contraste del chi cuadrado que los controles parentales no dependen del área residencial.

TABLA 14.
Tabla de frecuencias esperadas obtenida unas ves ya aplicada la formula a cada celda.

EDAD/AREA	URBANA	RURAL	TOTAL
13-16	41.9884	7.01155	49
16-19	167.8106	28.1893	197
19-22	242.5046	40.4953	283
22-25	211.6560	35.3439	247
25-28	210.7991	35.2008	246
28-31	192.8040	32.1959	225
31-34	157.6708	26.3291	184
34-37	145.6741	24.3258	170
37-40	94.2597	15.7402	110
40-43	58.2696	9.7303	68
43-46	29.9917	5.0082	35
46-49	2.57072	0.4292	3
TOTAL	1557	260	1817

Tomando en cuenta la tabla anterior, después de aplicar la frecuencia esperada a cada uno de los datos podemos deducir que los resultados siguen siendo exactos.

H₀: La variable edad de la madre es independiente al área residencial.

H₁: La variable edad de la madre si depende del área residencial.

Aplicamos las componentes del chi-cuadrado de la formula (11).

$$x^2 = \sum_{ij} \frac{(46 - 41.9884)^2}{41.9884} = 0.383$$

Luego se aplicó el estadístico de contraste de la formula (12).

$$X_{exp}^2 = \sum_{i=1}^k \sum_{j=1}^m \frac{(46 - 41.9884)^2}{41.9884} = 0.383 + \dots = 8.5719$$

Calculamos el valor Teórico de la fórmula (6) pero antes resolvemos la formula (1)

Para hallar el valor de los grados de libertad. En la tabla se puede notar un valor de 0 en el intervalo de 46-49 años debido a que este es menos del 20 % será pasable para el cálculo del chi-cuadrado.

$$(fil - 1) * (col - 1) = (12 - 1) * (2 - 1) = 11 * 1 = 11$$

```

Pearson's Chi-squared test

data: Tabla_Edad
X-squared = 8.5719, df = 11, p-value = 0.6613
    
```

FIGURA 9.
Valor de chi-cuadrado de las variables edad de la madre y el área residencial.

$$X_{0.05,11}^2 = 18.30$$

La decisión es: $X_{exp}^2 > X_{0.05,11}^2 = 8.5719 < 18.30$ se acepta H_0 como hipótesis nula por ende la edad de la madre no depende del área residencial urbana y rural.

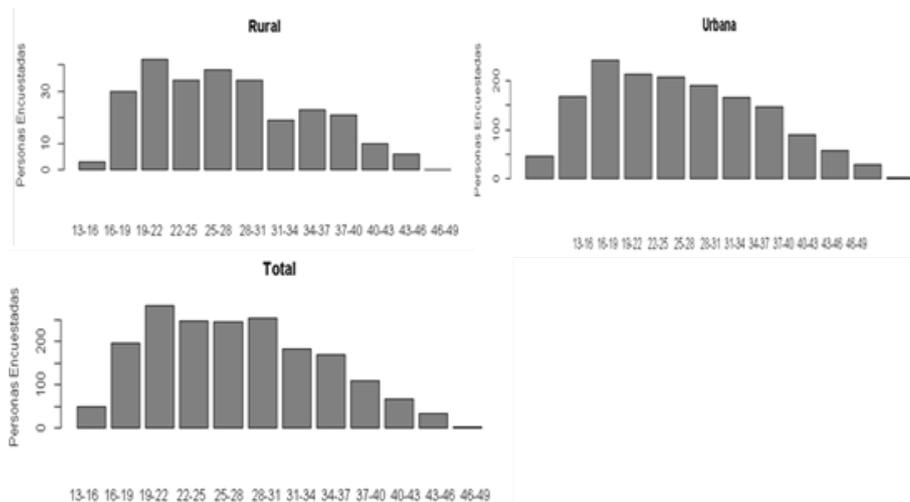


FIGURA 10.
Grafica realizada en Rstudio, muestra la tabla de contingencia entre las edades de las madres y el área residencial.

En la presente Grafica se pueden observar la frecuencia de pérdida materna en relación con la edad de la madre teniendo una mayor frecuencia en zonas urbanas en un intervalo de 19 a 22 años.

CONCLUSIONES

A través de este trabajo se obtuvieron varias pistas en el análisis de los resultados, lo que salta a la vista es que la mayoría de los casos de muerte fetal se presentan en la región litoral, principalmente en las zonas urbanas del Guayas (1558) de los casos registrados, siendo las madres en un intervalo de 13 a 49 años, más afectadas las madres de edad entre 19-22 años con 241 decesos fetales registrados.

Tomando en cuenta la relación entre la asistencia médica y la zona en donde vive, podemos concluir que el área rural tiene poca frecuencia en realizar estos seguimientos, esto se debe principalmente a la falta de información sobre cómo llevar un embarazo correctamente, especialmente si se lo planea llevar desde temprana edad, por otro lado en la zona urbana en las provincias mayores, hay factores que repercuten en la salud de la madre por consecuencia afectar al feto, debido a que este entorno se encuentra llena de contaminantes y estrés, lo cual provocaría la hipoxia en el feto siendo una de las mayores causas con 822 casos, a diferencia de las madres de la zona rural cuya pérdida puede ser ocasionada por el déficit de conocimiento sobre estos temas y los bajos fondos en general, ya que al tener escasez de economía no pueden permitirse chequeos médicos seguidos.

Finalmente se puede observar que el número de chequeos es más frecuente en cantidades menores a 8 (tanto en zonas urbanas como en rurales), debido que no hay relación frente a el área residencial, esto se puede deber a 2 motivos, quizás no sea accesible como ya se explicó, o quizás debido a las indicaciones seguidas por el galeno.

REFERENCIAS BIBLIOGRÁFICAS

- Aguilar, M. M. S., & Alvarado, J. O. R. (2000). Estadística y comportamiento organizacional. *Política y Cultura*, 13, 113-135.
- Bernal Cortés, D. P., & Cardona Rivas, D. (2014). CHARACTERIZATION OF PERINATAL MORTALITY IN MANIZALES, COLOMBIA, 2009-2012. *Hacia la promoción de la salud*, 19(2), 66-80.
- Carrillo, R. E., Arias, J. E., Mejillones, E. L., Naranjo, P. A., Torres, L. C., & de Sistemas Computacionales, C. (s. f.). Aplicación de tablas de contingencia para relacionar las variables área urbana y edad de la madre como factores que inciden en la mortalidad fetal.
- Cevallos-Torres, L., & Botto-Tobar, M. (2019a). Case study: Logistical behavior in the use of urban transport using the monte carlo simulation method. En *Problem-Based Learning: A Didactic Strategy in the Teaching of System Simulation* (pp. 97-110). Springer.
- Cevallos-Torres, L., & Botto-Tobar, M. (2019b). The system simulation and their learning processes. En *Problem-Based Learning: A Didactic Strategy in the Teaching of System Simulation* (pp. 1-11). Springer.
- CEVALLOS TORRES, L. J., VALENCIA MARTINEZ, N. A., & BARROS MORALES, R. L. (2017a). Análisis Estadístico Univariado. grupo Compás-Universidad de Guayaquil.
- CEVALLOS TORRES, L. J., VALENCIA MARTINEZ, N. A., & BARROS MORALES, R. L. (2017b). Análisis Estadístico Univariado. grupo Compás-Universidad de Guayaquil.
- De La Cruz-Oré, J. L. (2013). ¿Qué significan los grados de libertad? *Revista peruana de epidemiología*, 17(2), 1-6.
- Gómez, O. T. (2008). Una aplicación de la prueba chi cuadrado con SPSS. *Industrial data*, 11(1), 73-77.
- Gómez, P. S., & Soria, D. S. (2016). Concentración, dispersión y características sociodemográficas en la incorporación espacial de la migración peruana en la Ciudad de Córdoba, Argentina. *Población y Salud en Mesoamérica*, 14(1), 108-131.
- González-Galbán, H., & Herrera-León, L. I. (2015). Mortalidad infantil y preescolar en el estado de Baja California: Análisis de condicionantes biodemográficos relacionados con la historia reproductiva de la madre. *Población y Salud en Mesoamérica*, 12(2), 28-59.

- González, F. de J. G. (2014). Herramientas de calidad y el trabajo en equipo para disminuir la reprobación escolar. *Conciencia tecnológica*, 48, 17-24.
- Hernández, J., Hernández, M. G., Herrera, M., López, M., Rodríguez, J. G., & Guerrero, A. (2013). Morbimortalidad por síndrome de broncoaspiración meconial en recién nacidos de un hospital público del estado Cojedes: Venezuela. *Comunidad y salud*, 11(2), 11-17.
- Inzuna Cazares, S., & Jiménez Ramírez, J. V. (2013). Caracterización del razonamiento estadístico de estudiantes universitarios acerca de las pruebas de hipótesis. *Revista latinoamericana de investigación en matemática educativa*, 16(2), 179-211.
- Keck, M. E., & Sikkink, K. (2014). *Activists beyond borders*. Cornell University Press.
- Torres, L. C., Rodríguez, A. G., Cáceres, J. M. A., Veloz, G. S. D., Rivera, M. K. B., & Flores, R. M. A. (2018). Análisis estadístico de correlación entre las dosis de eritropoyetina y el nivel de hemoglobina en pacientes con insuficiencia renal crónica. *Revista de la Facultad de Ciencias Médicas de la Universidad de Guayaquil*, 19(1).
- ULLAURI, J. G. L., SOTO, J. M. O., & FERNÁNDEZ, M. D. J. V. R. (s. f.). FRECUENCIA Y CARACTERÍSTICAS DEL ÓBITO FETAL EN EL HOSPITAL VICENTE CORRAL MOSCOSO DE LA CIUDAD DE CUENCA, 2010 A 2014.
- Valencia-Nunez, E. R., Melendez-Tamano, C. F., Valle-Alvarez, A. T., Paredes-Salinas, J. G., Salinas, C. F. P., & Cevallos-Torres, L. J. (2018). Virtual classrooms and their use, measured with a statistical technique: The case of the Technical University of Ambato—Ecuador. 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), 1-6.
- Valladares, Z., García, V., Buján, V., Couceiro, E., & López, C. (2013). Muerte fetal intrauterina: ¿podemos actuar en su prevención? *Revista chilena de obstetricia y ginecología*, 78(6), 413-418.
- Zapata, C. M., Villegas, S. M., & Arango, F. (2006). Reglas de consistencia entre modelos de requisitos de UN-Metodo. *Revista Universidad Eafit*, 42(141), 40-59.

NOTAS

- [1] Investigador Asociado. E-mail: ronnie.carrillod@gmail.com
- [2] Investigador Asociado. E-mail: jose.ariasma@gmail.com
- [3] Investigador Asociado. E-mail: erick.limonm@gmail.com
- [4] Consultor Independiente. E-mail: pedro.naranjoi@gmail.com