

TECNOLÓGICO DE COSTA RICA

ESCUELA DE BIOLOGÍA

DESARROLLO DE UNA APLICACIÓN BIOINFORMÁTICA MEDIANTE R-SHINY PARA EL MONITOREO Y ANÁLISIS BIOESTADÍSTICO DE CASOS DE INTOXICACIÓN REPORTADOS AL CENTRO NACIONAL DE CONTROL DE INTOXICACIONES DE COSTA RICA.

Informe presentado a la Escuela de Biología del Tecnológico de Costa Rica como requisito parcial para optar al título de Licenciada en Ingeniería en Biotecnología

Lesli Natasha Mora Fonseca

Cartago, Junio, 2019



TECNOLÓGICO DE COSTA RICA

ESCUELA DE BIOLOGÍA

DESARROLLO DE UNA APLICACIÓN BIOINFORMÁTICA MEDIANTE R-SHINY PARA EL MONITOREO Y ANÁLISIS BIOESTADÍSTICO DE CASOS DE INTOXICACIÓN REPORTADOS AL CENTRO NACIONAL DE CONTROL DE INTOXICACIONES DE COSTA RICA.

Informe presentado a la Escuela de Biología del Tecnológico de Costa Rica como requisito parcial para optar al título de Licenciatura en Ingeniería en Biotecnología

Lesli Natasha Mora Fonseca

Cartago, Junio, 2019



RESUMEN

Día a día se genera gran cantidad de datos. Los cuales mediante el uso de herramientas bioinformáticas y estadísticas pueden convertirse en información para la toma de decisiones. Este trabajo pretende desarrollar una herramienta bioinformática para visualizar y analizar dinámicamente los datos generados por el Centro Nacional de Control de Intoxicaciones de Costa Rica mediante estadísticas específicas: el gráfico de embudo y la regresión de Poisson. Lo anterior facilitará al personal del centro la planificación de campañas u otras estrategias dirigidas a la disminución de las intoxicaciones en Costa Rica. Los resultados pueden variar según la definición del usuario. Sin embargo, el gráfico de embudo puede mostrar cantones en los que se presentan tasas de intoxicación fuera de los límites de control. La regresión de Poisson por otro lado, puede mostrar variables que están relacionadas al aumento o disminución de los números de intoxicaciones reportados cada año. Además se presentan estadísticas para determinar si el modelo generado es válido, si debe de ser modificado o descartado. Finalmente, se logró crear una interfaz web, con las estadísticas deseadas, que facilitará la toma de decisiones en el Centro Nacional de Control de Intoxicaciones de Costa Rica de acuerdo a un análisis dinámico de los datos sobre intoxicaciones en Costa Rica.

PALABRAS CLAVE: R, Shiny, Tóxico, Salud, Incidencia.

SUMMARY

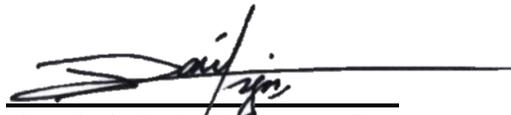
Daily, huge amounts of data are generated. By using bioinformatics tools and statistics those could be transformed into valuable information useful for decision making. This project pretends to develop a bioinformatics application to dynamically visualize and analyze data from *Centro Nacional de Control de Intoxicaciones* from Costa Rica with specific statistic methods: the funnel plot and Poisson regression. The previous, will facilitate the campaign and strategy planning to reduce the intoxications in Costa Rica. The results may vary depending of the user definition of variables. Nevertheless, the funnel plot can show the cities in which the intoxication rate is beyond the confidence intervals. Poisson regression, by the other hand, can help identify variables that are related with the increase or decrease in the number of intoxication cases reported every year. Statistics to determine the validity of the Poisson model are also available to consider if the model has to be adjusted or discarded. Finally, the creation of a web interface was achieved, with the desired statistical tests, which will simplify the decision making for the *Centro Nacional de Control de Intoxicaciones* based on a dynamic analysis on the intoxications in Costa Rica.

KEY WORDS: R, Shiny, Toxic, Health, Incidence.

DESARROLLO DE UNA APLICACIÓN BIOINFORMÁTICA MEDIANTE R-SHINY PARA EL MONITOREO Y ANÁLISIS BIOESTADÍSTICO DE CASOS DE INTOXICACIÓN REPORTADOS AL CENTRO NACIONAL DE CONTROL DE INTOXICACIONES DE COSTA RICA.

Informe presentado a la Escuela de Biología del Tecnológico de Costa Rica como requisito parcial para optar al título de Licenciatura en Ingeniería en Biotecnología

Miembros del Tribunal



**Ing. Raúl Trejos Espinoza MSc,
MBA**

Profesor Asesor



**Lic. Ana Gisselle Zeledón
Solano**

Asesora institucional



**Ing. Edgardo Gutiérrez Vega
Lector**

DEDICATORIA

A mis familiares y amigos
por su apoyo durante la
realización de este
trabajo.
Natasha Mora.

AGRADECIMIENTOS

A Edgardo Gutiérrez por compartir sus ideas de colaboración con el Centro Nacional de Control de Intoxicaciones de Costa Rica y así dar pie al desarrollo de esta tesis. Además por la constante guía en el desarrollo de este proyecto.

Al Centro Nacional de Control de Intoxicaciones de Costa Rica por abrirme las puertas y darme la oportunidad de desarrollar este trabajo. A la Dra. Ana Elvira Salas quien fue el contacto inicial en el centro y me ayudó a comprender las necesidades del mismo. A la Dra. Gisselle Zeledón por tener toda la voluntad y disposición para ayudar en lo que fuera necesario y brindarme su tiempo para el desarrollo del proyecto.

A Raúl Trejos por impulsarme a la finalización de esta etapa profesional.

A todos mis amigos y familiares por su apoyo e impulso.

ÍNDICE GENERAL

RESUMEN	i
DEDICATORIA	iii
AGRADECIMIENTOS	iv
ÍNDICE GENERAL	v
ÍNDICE DE CUADROS	vi
ÍNDICE DE FIGURAS.....	vii
ÍNDICE DE ANEXOS.....	vii
Introducción.....	1
Revisión de literatura.....	3
1. Intoxicaciones en Costa Rica y el mundo.....	3
2. Gráfico de embudo	4
3. Método de conteo: Regresión de Poisson	5
Objetivos	9
Materiales y Métodos	10
Resultados.....	15
<i>Gráfico de Embudo</i>	15
<i>Regresión de Poisson</i>	16
<i>Interactividad de la aplicación</i>	20
Discusión de Resultados	22
<i>Gráfico de Embudo</i>	22
<i>Regresión de Poisson</i>	24
<i>Interactividad de la aplicación</i>	27
Conclusiones y Recomendaciones	28
Bibliografía	29
ANEXOS	37

ÍNDICE DE CUADROS

Núm.	Título	Pág.
1	Resultados de la regresión de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.	17
2	Resultados de la Prueba de Sobredispersión para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.	17
3	Resultados del factor de inflación de la varianza generalizado para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango 0-3 años.	18
4	Resultados de la regresión de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.	19
5	Resultados de la Prueba de Sobredispersión para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.	19
6	Resultados del factor de inflación de la varianza generalizado para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.	20

ÍNDICE DE FIGURAS

Núm.	Título	Pág.
1	Gráfico de Embudo: Intoxicaciones en Costa Rica para el 2017 y censo 2011.	16
2	Gráfico de Embudo. Intoxicaciones en Costa Rica para el 2017 y censo 2011. Limitado para la provincia de San José.	16
3	Resultado del AIC para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.	18
4	Resultado del R^2 de McFadden para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.	18
5	Resultado del AIC para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.	19
6	Resultado del R^2 de McFadden para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.	20
7	Sección: “Actualizar Datos” de la interfaz web.	20
8	Filtros en sección: “Funnel Plot” de la interfaz web.	21
9	Sección: “Regresión de Poisson” de la interfaz web. Escenario I.	21
10	Sección: “Regresión de Poisson” de la interfaz web. Escenario II.	22

ÍNDICE DE ANEXOS

Núm.	Título	Pág.
1	Sección: Actualizar Datos de la interfaz web.	37
2	Sección: Funnel Plot de la interfaz web.	38
3	Sección: Regresión de Poisson, sub sección: Configuración de la interfaz web.	39
4	Sección: Regresión de Poisson, sub sección: Resultados de la interfaz web.	40
5	Código para la aplicación creada para el análisis de intoxicaciones en Costa Rica.	41

Introducción

Diariamente se generan grandes cantidades de datos en el mundo ya que de acuerdo con Cordeiro, Faloutsos y Traina, (2013) la información recolectada de manera digital en diversas áreas ha ido en aumento, no solo en cantidad, sino que también en complejidad. Sin embargo, los datos no tienen ningún valor a menos de que sean debidamente estructurados y analizados (Morreale et al., 2015).

La bioestadística, conocida como la aplicación de principios estadísticos en medicina, salud pública o biología (Sullivan, 2011) y la bioinformática, ciencia que involucra los planteamientos experimentales, los enfoques metodológicos y tecnológicos de la computación y la informática con el fin de facilitar el entendimiento del conocimiento científico (Capel y Yuste, 2016), son ramas de la biotecnología claves para el análisis de datos científicos.

Con el fin de responder preguntas y tomar decisiones, el análisis de los datos suele comenzar con métodos exploratorios de los datos y posteriormente análisis estadísticos (Friendly and Meyer, 2016). La ciencia de datos es justamente la que permite lo anterior, es el campo que procura extraer información valiosa de los datos mediante distintas etapas como: planteamiento de hipótesis, extracción de datos, exploración, manipulación de datos y comunicación de los resultados (Ganguly, 2017). Lo anterior luego de la detección de valores atípicos, grupos inusuales, tendencias, patrones y otros (Unwin, 2016).

Una de las herramientas informáticas que agilizan el análisis de los datos es el software especializado. Ejemplos de ellos son IBM SPSS, Matlab, Excel, Minitab, STATA, R, entre otros (Velázquez-Perea, 2017). Así mismo, el desarrollo de software o programación en distintos lenguajes como R y otros tales como Python (Vo et al., 2017) y Java (Chekanov, 2016) permiten un análisis eficaz.

Como lenguaje de programación, R fue creado a inicio de los años 90 por Ross Ihaka y Robert Gentleman y su primera versión fue liberada el 29 de Febrero del 2000 (Chambers, 2017). Se basó en el lenguaje S creado en los años 70 (Cotton, 2013). R es un software gratuito, libre y de código abierto, es decir, puede ser transformado por la comunidad (Kleinman y Horton, 2014).

R es un lenguaje de programación especializado en estadística (Gardener, 2012 y Hector, 2017). Además, es una excelente herramienta para la presentación de resultados de manera comprensible, mediante gráficos (Schmuller, 2016). Este lenguaje permite ejecutar todas las tareas requeridas para el análisis de datos: homogenizar, transformar, modelar, visualizar y aplicar diversos análisis estadísticos (Ganguly, 2017). Debido a la especialización de R para el análisis de datos y por la calidad de sus gráficos es que se seleccionó para cumplir con los objetivos de este trabajo.

Una variedad de bibliotecas, también conocidas como paquetes, en R, facilitan el acceder, visualizar y analizar datos (Charpentier, 2014). Los paquetes son la unidad fundamental para compartir código con funciones específicas. Un paquete puede incluir, datos, código y documentación para facilitar su uso y comprensión (Teutonico, 2013). Hasta Enero del 2015 se tenían unos 6000 paquetes en CRAN (Comprehensive R Archive Network por sus siglas en inglés), el cual es el principal repositorio donde se encuentran los paquetes de R (Wickham, 2015), cuyo número crece constantemente (Lim y Tjhi, 2015).

Shiny, es uno de los tantos paquetes desarrollados para R. Este integra R con código de HTML y JavaScript sin requerir conocimiento en los dos últimos, lo que facilita la generación de una aplicación web en las que el usuario interactúa con los datos sin requerir conocimientos en programación (Beeley, 2013). Gracias a lo anterior, se pueden realizar estadísticas complejas sin que el usuario lo note y desplegar el resultado en una amigable interfaz web (Resnizky, 2015).

Ejemplo de los datos que pueden ser analizados son los datos toxicológicos, esto gracias al uso de la bioinformática y a las herramientas previamente explicadas. La toxicología es la ciencia que investiga y comprende los agentes tóxicos (Hernández et al., 2010). Esta ciencia busca valorar, comprender, diagnosticar, y tratar los efectos adversos sobre la salud, que causan dichos agentes, directa o indirectamente a corto, mediano y largo plazo, sobre los seres vivos y sus poblaciones (Guitart, 2014).

Un agente tóxico, es toda sustancia que al entrar en contacto con el organismo por inhalación, ingesta o absorción a través de la piel, puede causar daño, enfermedad o hasta la muerte del organismo (Ghiggi, 2014). El impacto de un tóxico está determinado por varios factores, uno de los más determinantes es la dosis, seguido por circunstancias individuales y ambientales. Sin embargo, según Guitart, (2014), factores como: frecuencia, hora, ruta, velocidad de administración, especie, raza, genética, enfermedades pre-existentes, sexo, embarazo, lactancia, edad, alimentación, fármacos, sistema inmune o ambiente, pueden influir en la reacción.

Los datos disponibles para este proyecto son producto de la toxicología médica, lo que según Guitart, (2014), incluye el trabajo con individuos expuestos a un agente tóxico. En Costa Rica, los datos de intoxicaciones son recolectados por el Centro Nacional de Control de Intoxicaciones (CNCI) (Arroyo, Salas y Arias, 2014), tanto por consultas directas al CNCI o mediante el 911, por parte de personal médico y personas afectadas o allegados. El análisis de los datos, mediante distintas estadísticas, puede permitir transformar los datos recolectados en información esencial para la toma de futuras decisiones (Joyanes, 2016).

La estadística descriptiva es una herramienta que permite transformar datos en información mediante el uso de métodos gráficos, tabulares y numéricos (Vilar, 2005) facilitando así la descripción, visualización y resumen de los datos (Velázquez-Perea, 2017).

Además, se puede utilizar estadísticas específicas para responder a ciertas interrogantes. Como por ejemplo, ¿Existen cantones en el país con un comportamiento diferenciado en cuanto al número de intoxicaciones reportadas? La respuesta podría encontrarse con un gráfico de control estadístico, como lo es el gráfico de embudo o 'funnel plot' (Bottle y Aylin, 2016).

El gráfico de embudo permite, mediante la presentación visual de la media y los límites de control de los datos, resaltar los cantones en los que la tasa de intoxicaciones sobrepasa los valores esperados según límites establecidos (Ward et al., 2012). El gráfico es básicamente la comparación de una variable en varios grupos y se consigue al graficar el indicador de interés con respecto al tamaño de la muestra (Flowers, 2007), los cuales en este caso, corresponden al total de intoxicaciones por cantón dividido entre la población total del cantón y el total de personas por cantón, respectivamente.

Los límites de control, de acuerdo al tipo de datos de estudios pueden ser calculados de múltiples maneras (Flowers, 2007 y Szklo y Nieto, 2003). Para estudios en los que se utilizan conteo o tasas, según Eayres, (2010), es recomendable el uso del método de aproximación de Byar, el cual tiene como una de sus ventajas la facilidad de implementación.

Infinidad de estadísticas pueden ser utilizadas para responder otras preguntas de interés. Definir las variables que afectan el número de intoxicaciones reportadas en un periodo podría ser una de ellas. Al conocer esas variables podrían determinarse planes de acción para disminuir el número de casos reportados, pero más importante aún, el número de personas afectadas. Lo anterior se puede determinar mediante el uso de modelos de conteo como la Regresión de Poisson (Winkelmann, 2013).

Poisson no es el único modelo de conteo, sin embargo es el estándar (Cabo, 2014). La regresión de Poisson parte del supuesto de que la media y la varianza son iguales (Rodríguez, 2007). Además, implica el uso de una variable de respuesta (o dependiente), representada por un conteo o tasa, en este caso, del número de intoxicaciones o la tasa de intoxicaciones por periodo y una o más variables predictoras (o independientes) las cuáles pueden ser continuas y categóricas (PennState, 2018), como el Sexo, Provincia, Cantón, Edad, entre otros.

Facilitar el análisis de los datos recolectados por el Centro Nacional de Control de Intoxicaciones de Costa Rica va a permitir la comprensión y obtención de valiosa información del tesoro que de momento yace inexplorado. Los profesionales de la salud podrán analizar, interpretar y tomar acciones a partir de los distintos descubrimientos. Además, las conclusiones que obtengan podrán ser publicadas en artículos que promuevan el conocimiento entre los profesionales de la salud y ayuden a mejorar la imagen de Costa Rica como potencial líder en el control de intoxicaciones.

Finalmente, este trabajo tiene como objetivo desarrollar una herramienta bioinformática para visualizar y analizar dinámicamente los datos generados por el Centro Nacional de Control de Intoxicaciones de Costa Rica, mediante estadísticas específicas elaboradas con R y Shiny.

Revisión de literatura

1. Intoxicaciones en Costa Rica y el mundo.

Según un estudio de Arroyo et al., (2014), el agente tóxico responsable de la mayor cantidad de intoxicaciones en Costa Rica está representado por los medicamentos, seguido de los plaguicidas y productos de limpieza. En otros países como España, se tiene una situación similar en la que los medicamentos también encabezan la lista, solamente que en este caso son seguidos por los productos domésticos y el etanol (de la Torre, 2014).

En Europa, las intoxicaciones representan la quinta causa de muerte (Mintegi et al., 2015). En Costa Rica, las intoxicaciones accidentales representaron la cuarta causa de muerte del 2011 (Arroyo et al., 2014).

La Organización Mundial de la Salud (OMS) estima que entre 500,000 y 1 millón de personas se intoxican con plaguicidas, accidental o intencionalmente (Martínez-Valenzuela y Gómez-Arroyo, 2007). De estas, pueden llegar a morir hasta 20,000 anualmente (Guzmán-Plazola et al., 2016).

Es debido a los números anteriores y a la importancia de cada uno de los habitantes que se deben tomar medidas. Se debe analizar detalladamente si las intoxicaciones de nuestro país se centran en una región, grupo etario, sexo o causa. De acuerdo con de la Torre, (2014) la mayoría de las intoxicaciones son accidentales, voluntarias o por errores de dosificación.

En España, la mayoría de las intoxicaciones accidentales se da en niños por ingesta de medicamentos o productos del hogar que se encuentran almacenados en recipientes distintos al original (Mintegi et al., 2015). En otros casos, las intoxicaciones se pueden dar por medicamentos usuales a los cuales las personas les han tomado demasiada confianza y han obviado la existencia de posibles efectos adversos, en caso de abuso, un ejemplo es la acetaminofén, el cual puede llegar a causar una serie de síntomas incluyendo hepatotoxicidad (Vargas, 2016).

En otros casos, las intoxicaciones no solo se dan por ingesta, accidental o intencional, se pueden dar por exposición prolongada, como es el caso de los plaguicidas. Estos son los responsables del 99% de las intoxicaciones agudas en actividades agrícolas debido a falta de educación sobre las medidas preventivas que se deben tomar al utilizarlos, el mal almacenamiento y la mala disposición de los envases (Guzmán-Plazola et al., 2016). Según, Vargas, (2015), un ejemplo claro de lo anterior en nuestro país ocurre en la provincia de Guanacaste. El plaguicida de mayor uso en el país, involucrado en intoxicaciones por accidentes laborales, pero principalmente por ingesta accidental o suicida es el “paracuat” o también conocido como Gramoxone® (Viales, 2014).

2. Gráfico de embudo

Parte del análisis de los datos de intoxicaciones en Costa Rica puede realizarse mediante un gráfico de embudo o también conocido como “funnel plot”, el cual es un gráfico para el control estadístico (Bottle y Aylin, 2016).

El gráfico de embudo compara el indicador de interés (total de intoxicaciones por cantón dividido por la población total del cantón) con respecto al tamaño de la muestra (total de personas por cantón) (Flowers, 2007 y Fox, Negrete-Yankelevich y Sosa, 2015). El gráfico, gracias a la presentación visual de los datos, media y límites de control, resalta los cantones en los que la tasa de intoxicaciones sobrepasa los valores esperados según los límites establecidos (Ward et al., 2012).

Los límites de control se pueden determinar de distintas maneras de acuerdo a los tipos de datos que se estudian (Szklo y Nieto, 2003 y Flowers, 2007). Para el caso específico de conteos o tasas se aconseja, según Eayres, (2010), utilizar el método de aproximación de Byar, considerando como una de sus ventajas la facilidad de cálculo. El método de aproximación de Byar, utiliza las siguientes fórmulas:

Límite inferior:

$$r_{inferior} = \frac{O_{inferior}}{n} \text{ donde,}$$
$$O_{inferior} = O \times \left(1 - \frac{1}{90} - \frac{z}{3\sqrt{O}}\right)^3$$

Límite superior:

$$r_{superior} = \frac{O_{superior}}{n} \text{ donde,}$$

$$O_{superior} = (O + 1) x \left(1 - \frac{1}{9(O + 1)} - \frac{z}{3\sqrt{(O + 1)}} \right)^3$$

Donde,

- **O** representa la tasa de intoxicaciones definida como el número de intoxicaciones/población total.
- **z** representa el valor del percentil $100(1 - \alpha/2)$ de una distribución normal, donde α para una confianza del 95% será de 0,05.
- **n** representa el tamaño de la muestra.

Fuente: Eayres, (2010).

Son los límites antes mencionados los que permiten evidenciar comportamientos anormales que indican potenciales problemas mas no proveen una razón clara para tal comportamiento (Morton et al., 2013). Sin embargo, queda a interpretación del analista las medidas que se quieran tomar, en el caso de un número de intoxicaciones por cantón anormalmente elevado, podría facilitar al analista la identificación de los sitios del país en los que según sugiere Flowers, (2007) se podría requerir investigación adicional.

Ejemplo del uso de los gráficos de embudo es la implementación de estos para monitorear qué instituciones requieren estudios más rigurosos en auditorías (Morton et al., 2013). Más relacionado a ciencias de la salud, se puede determinar el éxito de los procedimientos quirúrgicos según el tipo de procedimiento o institución en la que se realizan los procedimientos, así como la tasa de mortalidad en niños (Flowers, 2007, Morton et al., 2013 y Ieva y Paganoni, 2014) y en este caso, la tasa de intoxicaciones en los distintos cantones de nuestro país.

3. Método de conteo: Regresión de Poisson

Cuando se habla de intoxicaciones y de la cantidad de variables que pueden afectar la incidencia de estas (de la Torre, 2014 y Guzmán-Plazola et al., 2016), así como la severidad (Guitart, 2014), resulta importante saber qué factores tienen mayor incidencia en el número de intoxicaciones por año en Costa Rica.

Lo anterior se puede determinar con un modelo de conteo. De estos, el más utilizado según Correa y Barrera, (2018), es el modelo de Poisson, el cual de acuerdo con Cabo, (2014), es considerado el estándar. Así mismo, el modelo o regresión de Poisson es recomendado para analizar datos de conteos (Correa y Barrera, 2018) con una distribución de Poisson (Ryan, 2013), y su asociación con múltiples variables (Hayat y Higgins, 2014).

La distribución de Poisson es un modelo lineal generalizado (Hayat y Higgins, 2014) que asume que los sucesos son independientes uno de otro y que la probabilidad de que ocurran son constantes en el tiempo y en el espacio (Zoni, 2018). Para realizar el cálculo de la regresión de Poisson se debe de contar con una variable dependiente o de respuesta, en este caso, el número de intoxicaciones por año y una o varias variables independientes o predictoras (PenState, 2018). Las últimas podrían ser continuas y categóricas (PenState, 2018), como el Sexo, Provincia, Cantón, Causa, Severidad, entre otros.

Esta distribución está definida por un solo parámetro lambda (λ), que representa tanto a la media como a la varianza, por lo tanto, podemos decir que parte del supuesto de que la media y la varianza son iguales (Hayat y Higgins, 2014 y Rodríguez, 2007).

El cálculo de la regresión de Poisson, utiliza la siguiente fórmula:

$$\log(y) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_nx_n \dots$$

Donde,

- y representa la variable de respuesta (dependiente), la que corresponde al número de casos reportados.
- $\beta_0, \beta_1, \dots, \beta_n$ son los coeficientes numéricos del modelo.
- x_1, x_2, \dots, x_n representa las n variables predictoras (independientes).

Fuente: PenState, (2018).

De acuerdo con Kaminski, (2004) y Zoni, (2018), los resultados pueden presentarse como tasas de incidencia o IRR (*Incidence Rate Ratios* por sus siglas en inglés), para los cuales se eleva e a los coeficientes estimados resultantes de la regresión de Poisson, $((e)^{\text{estimado}})$, de igual manera con los intervalos de confianza (Rabe-Hesketh y Skrondal, 2008).

Se tiene que si el valor del IRR es **menor a 1**, se considera que hay una menor cantidad de intoxicaciones con respecto al nivel de referencia. Según ejemplos de Meyers, Neuberger y He, (2009), el porcentaje de disminución de intoxicaciones se expresa porcentualmente como el resultado de **(1-IRR)*100**. Un valor de IRR igual a 1, indicaría que no hay disminución o aumento de las intoxicaciones con respecto al nivel de referencia (Harris, Yang y Hardin, 2012).

Por otro lado si el valor del IRR es **mayor a 1**, se considera que hay una mayor cantidad de intoxicaciones con respecto al nivel de referencia. Y de acuerdo con Harris et al., (2012) se puede expresar como porcentaje de aumento de intoxicaciones, calculando el resultado de **(IRR-1)*100**. Además, se puede expresar como 2, 3, 4, ..., n veces más intoxicaciones con respecto al nivel de referencia. Junto con la tasa de incidencia, debe de considerarse la significancia para cada nivel evaluado de acuerdo con las variables definidas (Shephard, 2011 y Marcus y Zgierska, 2012).

3.1 Verificación del modelo

A veces, los datos no cumplen los supuestos especificados para un modelo, por lo tanto se debe de considerar ajustar el modelo o probar con otros modelos (Hayat y Higgins, 2014). Para determinar si el modelo de Poisson aplicado es funcional, se pueden utilizar, en conjunto, una serie de fórmulas para comprobar el cumplimiento de supuestos (Hawkins y Olwell, 2012), comparar modelos (Kramer y Eden, 2016), determinar la eficacia del modelo (Westfall y Henning, 2013 y Fávero y Belfiore, 2019) o también la confiabilidad de los estimadores (Montgomery y Runger, 2018). Así mismo, tomando en cuenta ciertos factores, se pueden utilizar métodos para ajustar los errores (Salinas-Rodríguez, Manrique-Espinoza y Sosa-Rubí, 2009).

3.1.1 Prueba de sobredispersión:

Permite determinar si hay sobredispersión (varianza > media) o subdispersión (varianza < media) de los datos (Hayat y Higgins, 2014) y por tanto el no cumplimiento del supuesto de que la media es igual a varianza (Rodríguez, 2007). La prueba de sobredispersión se calcula de la siguiente manera:

$$VAR[\gamma] = \mu + \alpha * f(\mu)$$

Donde,

- \mathbf{y} representa la variable de respuesta (dependiente), la que corresponde al número de casos reportados.
- μ es la media de los datos.
- $f(\mu)$ es la función de transformación de la media, tal que $f(\mu) = \mu^k$, donde k es 1 o 2, produciendo así, formulaciones lineales o cuadráticas, respectivamente.
- α es un coeficiente estimado mediante una regresión lineal tal que $Var(\mathbf{y}) = -1 + f(\mu)$ se compara con el correspondiente estadístico t o z de la distribución normal estándar, bajo la hipótesis nula de que $\alpha = 0$.

Fuente: Friendly y Meyer, (2016) y Kleiber y Zeileis, (2018).

En caso de existir sobredispersión, es decir, valores de dispersión mayores a 0 y valor de p menor a 0.05 para un 95% de confianza (Hilbe, 2011), en al menos una de las dos pruebas se sugiere ajustar el modelo o utilizar el modelo de regresión binomial negativa (Hayat y Higgins, 2014), cuadrático o lineal (Fox, Weisberg y Price, 2018). Por otro lado, de presentarse un valor de dispersión menor a 0 y un valor de p menor a 0.05, habría sub-dispersión, lo cual podría corregirse con un modelo de Hurdle (Hilbe, 2014).

3.1.2 Criterio de información de Akaike o AIC

Es una medida para comparar modelos que toma en cuenta el ajuste de los datos y el número de variables que se utilizan (Kramer y Eden, 2016). Se define cómo:

$$AIC = -2\log L(\theta_{ML}) + 2\rho$$

Donde,

- θ_{ML} representa el parámetro estimado mediante máxima verosimilitud.
- $L(\theta_{ML})$ representa la función de verosimilitud de los datos para el parámetro seleccionado, estimado por máxima verosimilitud.
- ρ representa el número de parámetros en el modelo.

Fuente: Kramer y Eden, (2016).

De acuerdo con Kramer y Eden, (2016) y Westfall y Henning, (2013), entre menor AIC tenga un modelo, mejor es con respecto al otro u otros con los que se compara. Sin embargo, no deben de tomarse decisiones basadas únicamente en este criterio, sino en un conjunto (Kunert, 2016).

3.1.3 R^2 de McFadden

Dado a que no hay un R^2 para la regresión de Poisson se usa el Pseudo R^2 o R^2 de McFadden, el cual no debe de interpretarse de la misma manera que el R^2 común (Kunert, 2016). Es utilizado para

escoger entre modelos, siendo mejor el que presente un mayor valor de Pseudo R^2 , los cuales pueden ir de 0 a 1 (Fávero y Belfiore, 2019). El pseudo R^2 de McFadden se define como:

$$pseudo R^2 = \frac{-2LL_0 - (-2LL_{max})}{-2LL_0}$$

Donde,

- LL_0 representa la máxima verosimilitud logarítmica para el modelo cuando solo hay una constante α , conocido como el modelo nulo.
- LL_{max} representa la máxima verosimilitud logarítmica del modelo evaluado.

Fuente: Fávero y Belfiore, (2019).

El Pseudo R^2 de McFadden no representa la proporción de la variabilidad que es explicada por las variables independientes (Kunert, 2016). Little, (2013), indica que se refiere a cuan mejor que el modelo nulo (sin variables), el modelo evaluado explica la variabilidad de los datos. Un valor de 0 indica que el modelo evaluado explica tanto como el modelo nulo, es decir, que las variables predictoras no contribuyen a explicar la variabilidad de los datos. Por otro lado, un valor cercano a 1 indica que se explica cerca de un 100% más que el modelo nulo, lo cual indica que las variables predictoras tienen valor explicativo. Cabe resaltar que pese a que los valores permitidos son de 0 - 1 no es sorprendente ni incorrecto que los valores se mantengan bajos (Little, 2013).

3.1.4 Factor de la Inflación de la Varianza

Es un estadístico utilizado para detectar multicolinealidad, es decir, dependencia entre variables y así determinar si algunas deben de ser eliminadas de la regresión (Montgomery y Runger, 2018). El factor de la inflación de la varianza se calcula como:

$$FIV = \frac{1}{1 - R_j^2}$$

Donde,

- $\frac{1}{1 - R_j^2}$ representa la función de la correlación múltiple.
- R^2 representa el coeficiente de la función de correlación múltiple resultante de la regresión.
- j representa una de las variables elegidas para el modelo, expresado como la número j .

Fuente: Tattar, Ramaiah y Manjunath, (2016) y Montgomery y Runger, (2018).

Sin embargo, la función anterior no aplica cuando se tienen variables con múltiples grados de libertad, por lo tanto, se usa la versión generalizada del factor de inflación de la varianza, lo cual considera los grados de libertad de los predictores. La medida resultante, considerando los grados de libertad se conoce como $GFIV^{1/2GL}$, siendo GL los grados de libertad. Con el fin de utilizar las reglas que se tienen para los valores de VIF, que se explican más adelante, se puede elevar a la segunda potencia el $GFIV^{1/2GL}$, obteniendo así, $(GFIV^{1/2GL})^2$, que es lo mismo que decir $GFIV^{1/GL}$ (Fox y Monette, 1992 y Fox, 2002).

Cualquier valor de GFIV que supere el valor de 1 presenta algún grado de multicolinealidad (Montgomery y Runger, 2018). Se considera una regla que valores de GFIV mayores a 10 deben de ser

revisados y se sugiere eliminar la variable con el mayor GVIF superior a 10 ya que presentan multicolinealidad y re-evaluar el modelo (Tattar et al., 2016). Según Montgomery y Runger, (2018), algunos autores más rigurosos consideran que el valor de GFIV para las variables no debería de superar 4 o 5.

3.1.5 Errores Robustos

Los errores robustos se utilizan para compensar posibles violaciones del supuesto de que la media y la varianza son iguales en la regresión de Poisson (Cameron y Trivedi, 2009). Los mismos pueden ser calculados gracias al paquete de R, *sandwich* 2.5.0 (Cameron y Trivedi, 2009), en conjunto con otras funciones base de R (Universidad de California en Los Ángeles, 2019).

La función *vcovHC*, del paquete de R *sandwich* 2.5.0, puede utilizar una variedad de estimadores (Universidad de California en Los Ángeles, 2019). Según Girón, (2017), al ser el modelo de Poisson un modelo heterocedástico, con el estimador de White se pueden ajustar los errores para obtener errores robustos. La heterocedasticidad ocurre cuando la variabilidad de una variable dependiente no es constante a través de los distintos niveles de las variables predictoras (Libby y Thorne, 2017).

Es la regresión Poisson en conjunto con las estadísticas antes mencionadas que permitirán identificar variables importantes en el número de intoxicaciones en Costa Rica que se reportan por año. Según Weaver et al., (2015), la regresión de Poisson es frecuentemente usada para analizar datos de hospitalizaciones, cuando los resultados se expresan como conteos, por ejemplo, número de días en el hospital.

Existen muchos otros ejemplos. En el área epidemiológica la regresión de Poisson se ha usado para evaluar asociaciones entre la incidencia de colitis microscópica en un condado de los Estados Unidos y la edad y sexo de los pacientes (Gentile et al., 2014). Así mismo, se ha utilizado en investigaciones sobre la incidencia de diabetes tipo I en niños de Shanghai, China (Zhao et al., 2014). En Estados Unidos, se utilizó un modelo de Poisson para examinar la incidencia de yersiniosis y su relación con raza y grupo etario (Chakraborty et al., 2015). Son solo algunos ejemplos, pero fácilmente se podrían citar muchos más (Theisen, Drabik y Stock, 2012, daCruz et al., 2012, Staggs, 2013, Bergman et al., 2015, Watanitanon et al., 2018 y Alves et al., 2019).

Al igual que en los estudios antes mencionados, el personal del Centro Nacional de Control de Intoxicaciones de Costa Rica podrá identificar variables influyentes en el número de intoxicaciones en Costa Rica, información que les permitirá tomar acciones focalizadas, con el propósito de disminuir las intoxicaciones en nuestro país.

Objetivos

Objetivo General

Desarrollar una herramienta bioinformática para visualizar y analizar dinámicamente, mediante estadísticas específicas, los datos generados por el Centro Nacional de Control de Intoxicaciones de Costa Rica.

Objetivos Específicos

- Retornar un gráfico de embudo que permita identificar cuándo los reportes de intoxicaciones por cantón sobresalen con respecto a otros cantones.

- Facilitar la identificación de variables que afectan el número de intoxicaciones que se reportan por año, mediante el uso del modelo de Poisson.
- Proveer una interfaz de interacción construida con R-Shiny en la que se puedan analizar y visualizar los datos del Centro Nacional de Control de Intoxicaciones de Costa Rica mediante las estadísticas y gráficas definidas anteriormente.

Materiales y Métodos

Materiales

- R 3.5.0
- RStudio 1.1.447
- Bibliotecas de R y sus respectivas versiones
 - *shiny* 1.1.0
 - *shinydashboard* 0.7.0
 - *tidyr* 0.8.1
 - *readxl* 1.3.0
 - *ggplot2* 3.1.0
 - *scales* 1.0.0
 - *plotly* 4.8.0
 - *DT* 0.4
 - *sandwich* 2.5.0
 - *modEvA* 1.3.2
 - *car* 3.0.2
 - *AER* 1.2.6
 - *stats* 3.5.0
- Datos sobre intoxicaciones proporcionados por el Centro Nacional de Control de Intoxicaciones de Costa Rica, con formato similar (nombres de columnas iguales y niveles similares en el caso de las columnas con información sobre tipo de intoxicación, sexo y reportes mediante el 911) a los datos de 2017, 2018 y 2019.
- Datos sobre censo poblacional con información de población por cantón y provincia.

Métodos

Obtención de datos

Como referencia para la creación de la aplicación se utilizaron datos 2017 del Centro Nacional de Control de Intoxicaciones de Costa Rica los cuales prescinden de datos que permitan la identificación de los pacientes. Las columnas requeridas fueron:

- Cantón: indica los cantones de Costa Rica. Niveles como Alajuelita, Abangares.
- Provincia: indica las provincias de Costa Rica. Niveles como San José, Heredia.
- Fecha: fecha del reporte de intoxicación. Niveles indicando dd/mm/aaaa.
- Edadanos: edad de los pacientes. Es luego transformada para mostrar rangos etarios.
- Sexo: indica el sexo de los pacientes. Niveles; Femenino, Masculino y No Anotado.

- Toxico1: Indica el tóxico principal. Ej: 2,4-D, Abejas.
- Codtox1: agrupa los tóxicos en categorías Ej: Medicamentos, Plantas.
- Causa: indica la causa de la intoxicación. Ej: Accidente, Automedicación.
- Ruta: indica la ruta de acceso del tóxico al cuerpo del paciente. Ej: Ingestión, Piel.
- Severidad: Indica la severidad de la intoxicación. Niveles como Leve, Fatal.
- NueveOnce: indica si el reporte de la intoxicación se dio mediante el número 911. Niveles Si o No.
- NoConsulta: identificador único para cada reporte de intoxicación. Valores numéricos 1, 2, 3...
- Entidades: indica el sitio en el que se reporta la intoxicación. Ej: Casa y Alrededores, Clínica Abangares.
- Toxico2: indica el 2do tóxico de importancia si más de uno estuvo involucrado.
- Codtox2: indica la categoría del 2do tóxico de importancia.
- Toxico3: indica el 3ro tóxico de importancia si más de dos estuvieron involucrados.
- Codtox3: indica la categoría del 3ro tóxico de importancia.
- Toxico4: indica el 4to tóxico de importancia si más de tres estuvieron involucrados.
- Codtox4: indica la categoría del 4to tóxico de importancia.
- Toxico5: indica el 5to tóxico de importancia si más de cuatro estuvieron involucrados.
- Codtox5: indica la categoría del 5to tóxico de importancia.
- Toxico6: indica el 6to tóxico de importancia si más de cinco estuvieron involucrados.
- Codtox6: indica la categoría del 6to tóxico de importancia.
- TipoIntox: indica si la intoxicación es en humanos o animales.
- Intoxicación: indica si la llamada se debió a una intoxicación o una asesoría.

Además se solicitaron datos del censo poblacional del 2011 al Instituto Nacional de Estadística y Censo (INEC, 2011). Las columnas utilizadas fueron: Cantón, Provincia, Hombres, Mujeres y Total, las últimas tres indicando número de personas, ya sea según el sexo o total.

Limpieza y transformación de datos

Los datos del Censo 2011 se ajustaron en una hoja de Excel (Microsoft Office 365 ProPlus) que permitirá actualizar fácilmente los datos una vez que existan actualizaciones con nuevos censos.

Se utilizó el paquete *readxl* 1.3.0, con el objetivo de leer los datos de tanto el Censo 2011 como del Centro Nacional de Control de Intoxicaciones de Costa Rica, en formato *.xls o *.xlsx (Wickham y Bryan, 2019).

Se removieron los datos de animales y de asesorías ya que no son relevantes para el estudio. Se estandarizaron los valores correspondientes a la columna NueveOnce, para representar como “Si” o “No” si los reportes han sido o no mediante la línea 911. Se calculó una nueva columna RangoEtario de acuerdo con las especificaciones del CNCI (0-3, 4-6, 7-12, 13-15, 16-19, 20-29, 30-39, 40-49, 50-59, 60-69 años y >70 años.).

Se utilizó el paquete *tidyr* 0.8.1 para hacer transformaciones de los datos, como por ejemplo, separar el identificador asignado a los cantones o provincias en los datos del Centro Nacional de Control de Intoxicaciones de Costa Rica (Wickham y Henry, 2019).

Desarrollo de estadísticas de control

Se desarrolló, con código de R (versión 3.5.0), un gráfico de embudo en el que se graficará:

- En el eje Y: la tasa de intoxicaciones por cantón por cada 100,000 habitantes, calculada como el total de intoxicaciones por cantón dividido por la población total del cantón según el Instituto Nacional de Estadística y Censos (INEC, 2011).
- En el eje X: la población total de cada cantón.

Además, el gráfico muestra el promedio de la tasa de intoxicaciones y los límites de control, estos últimos, calculados mediante el método de aproximación de Byar utilizando las fórmulas descritas por Eayres, (2010).

Para la creación de la línea que representa la media, se calculó el total de intoxicaciones tomando en cuenta la tasa, lo cual se obtiene al multiplicar el promedio de la tasa (**m**) por una secuencia (**nseq**) de 1000 números inician desde el valor mínimo de la población de los cantones hasta el valor máximo de la población de los cantones.

Se calcularon los límites de control utilizando las fórmulas descritas por Eayres, (2010). Una vez obtenido el resultado de esas fórmulas, se dividió entre la secuencia de 1000 valores nombrada **nseq**, y así se consiguió los valores a utilizar para interpolar los límites de control de los datos.

Se creó un vector de color de acuerdo con los valores de tasa de intoxicaciones para cada cantón y su relación con los límites de control. Si estaban dentro de los límites, verde, entre el límite del 95% de confianza y el de 99% de confianza, amarillo y sobre el límite del 99% de confianza, rojo.

Finalmente, con los valores obtenidos, se utilizó el paquete *ggplot2* 3.0.1 el cual es utilizado para la creación de visualizaciones (Wickham et al., 2019). Además se hizo uso del paquete *scales* 1.0.0 el cual permitió presentar los valores del total de población como valores continuos sin el uso de la notación científica seleccionada por defecto al crear el gráfico (Wickham, 2018).

Desarrollo de un modelo de conteo

Se desarrolló, con código de R, un modelo de conteo utilizando la regresión de Poisson, la cual utiliza la fórmula descrita por PenState, (2018). Específicamente la función de R que se utilizó para obtener la regresión de Poisson fue la *glm* del paquete *stats* 3.5.0 (R Core Team, 2018).

Con el objetivo de calcular la regresión de Poisson, se definió como la variable de respuesta o dependiente el número de intoxicaciones y como las posibles variables predictoras, las columnas: Cantón, Provincia, RangoEtario, Sexo, Toxico1, Codtox1, Causa, Ruta, Severidad, NueveOnce y Entidades. Se permitió la selección de hasta tres variables predictoras y la posibilidad de escoger si se permiten interacciones entre ellas o no.

De acuerdo con las variables predictoras seleccionadas, se generó un conteo del número de intoxicaciones para cada combinación de los diferentes niveles de las variables seleccionadas. Usando los conteos anteriores y la función *glm* del paquete *stats* 3.5.0 (R Core Team, 2018), el modelo se definió como:

$glm = (Resp \sim var1 + var2 + var3, family = poisson, data = Datos)$

ó

$glm = (Resp \sim var1 * var2 * var3, family = poisson, data = Datos)$, para habilitar el modelo con interacciones.

Donde,

- **Resp**, es la variable de respuesta, el número de intoxicaciones.
- **var**, representa las hasta tres variables de respuesta que se pueden elegir, entre: Cantón, Provincia, RangoEtario, Sexo, Toxico1, Codtox1, Causa, Ruta, Severidad, NueveOnce y Entidades
- **Datos**, representa los datos siendo usados, con los conteos de acuerdo con las variables seleccionadas.
- **family = poisson**, se refiere a la familia del modelo lineal generalizado seleccionado.

Fuente: R Core Team, (2018).

Una vez desarrollado el modelo se presentaron los resultados como tasas de incidencia, para lo cual se elevó la constante matemática o número e a los coeficientes estimados de la regresión, de igual manera se elevó e a los coeficientes obtenidos para los intervalos de confianza Zoni, (2018). Además, para compensar posibles violaciones del supuesto de que la media y la varianza son iguales en la regresión de Poisson (Cameron y Trivedi, 2009), se calcularon los errores robustos, utilizando la función *vcovhc* y el estimador de White del paquete *sandwich* 2.5.0 (Zeileis y Lumley, 2019).

De acuerdo con recomendaciones de Sribney y Wiggins (2019) se derivó un estimado para los errores estándar de las tasas de incidencia, por lo tanto, se utilizó la siguiente fórmula:

$$Errores = e^y * stdErr$$

Donde,

- **e** representa la constante matemática, cuyo valor es aproximadamente 2.71828.
- **y** representa los coeficientes resultantes del modelo de Poisson.
- **stdErr** representa el valor de los errores robustos de los coeficientes del modelo de Poisson.

Fuente: Sribney y Wiggins, (2019).

Junto con la regresión, se calcularon una serie de estadísticas para comprobar el cumplimiento de supuestos, permitir la comparación de modelos, determinar la eficacia del modelo y la confiabilidad de los estimadores.

La prueba de sobredispersión, utilizado para detectar el incumplimiento del supuesto de que la media es igual a varianza (Rodríguez, 2007), se calculó la función *dispersiontest* del paquete *AER* 1.2.6 (Kleiber y Zeileis, 2018).

El criterio de información de Akaike o AIC, medida para comparar modelos que toma en cuenta el ajuste de los datos y el número de variables que se utilizan (Kramer y Eden, 2016), se extrajo del resumen del resultado del modelo de regresión de Poisson, usando funciones base de R.

El R^2 de McFadden, utilizado para comparar el modelo con el modelo nulo (Fávero y Belfiore, 2019), se determinó usando la función *RsqGLM* del paquete *modEVA* 1.3.2 (Barbosa et al., 2016).

El factor de la inflación de la varianza generalizado, utilizado para detectar multicolinealidad (Montgomery y Runger, 2018), se calculó con la función *vif* del paquete *car* 3.0.2 (Fox et al., 2018).

Desarrollo de aplicación para el análisis de los datos

Una vez desarrollado el código para el gráfico de embudo y la regresión de Poisson. Se creó una interfaz web interactiva mediante el uso de *shiny* 1.1.0 (Chang et al., 2019) y RStudio 1.1.447. Para dar formato a la interfaz se utilizó el paquete *shinydashboard* 0.7.0 (Chang y Borges, 2018). Con el fin de presentar el gráfico de embudo y permitir interacción con el mismo, seleccionando cantones específicos, se utilizó la función *ggplotly* del paquete *plotly* 4.8.0 (Sievert et al., 2019). De forma similar, para presentar las tablas de resultados con un formato más elegante que el usado por defecto, se utilizó la función *datatable* del paquete *DT* 0.4 (Xie, Cheng y Tan, 2019).

Mediante *shiny* 1.1.0 (Chang et al., 2019) se logró habilitar diferentes secciones para separar la carga de datos (I), el gráfico de embudo (II) y la regresión de Poisson (III). Para la primera (I), se incluyó la posibilidad de cargar los datos de tanto el CNCI como el INEC y un botón para actualizar datos.

En la segunda (II), se incluyó funcionalidad para permitir una serie de filtros (Provincia, Fecha de Consulta, Edad, Sexo, Tóxico1, CodTox1, Causa, Ruta, Severidad y NueveOnce) que limitaron los datos utilizados para generar el gráfico de embudo. Se habilitaron opciones para el reinicio de los filtros, colapsar tablas, resaltar con color negro cantones específicos, cambio interactivo en la cantidad de habitantes utilizados para expresar las tasas. Para el **Escenario I** uno no se aplicó filtros, para el **Escenario II**, se aplicó un filtro donde se limitaron los datos a la provincia de San José.

En la tercera (III), se incluyó una sección para la configuración de datos y otra para la visualización de los resultados de la regresión de Poisson. En la primera sub-sección se creó la opción para elegir hasta 3 variables predictoras. Se definieron valores de referencia por defecto (Sexo = Femenino, NueveOnce = No), de acuerdo al nivel con mayor número de intoxicaciones (aplica para las variables Provincia, RangoEtario y Severidad) o con respecto al nivel "Otros". "Otros" es calculado de acuerdo con el número especificado en la sección de configuración e incluye todos los niveles cuyo conteo de intoxicaciones sea menor al valor especificado (aplica para las variables Cantón, Tóxico1, CodTox1, Causa, Ruta y Entidades). Sin embargo, se creó la opción de cambiar los valores de referencia. Así mismo, se agregó funcionalidad que permitió cambiar de un modelo sin interacciones a uno con interacciones. Además, se generó la opción de filtros dinámicos, de acuerdo con la variables predictoras seleccionadas.

En el **Escenario I** de la regresión de Poisson, se seleccionó las variables Provincia y RangoEtario. Se limitó las provincias a Alajuela, Cartago, Heredia y San José, los rangos etarios se limitaron a aquellos

que incluyen edades de 19 años o menos. Los niveles de referencia fueron **San José** y **a. 0-3**, para Provincia y Rango Etario respectivamente. No se permitieron interacciones.

En el **Escenario II** de la regresión de Poisson, se seleccionó las variables Cantón y Tóxico1. Se limitó los cantones a San José, Heredia y Alajuela, los tóxicos no fueron limitados mediante filtros. Se definió el número de intoxicaciones para generar el nivel de “Otros” en las columnas predictoras seleccionadas, 0 para cantón, lo cual no genera nivel “Otros” y por lo que se definió como nivel de referencia San José y 80 para Tóxico1, de modo de que se agruparon todos los tóxicos cuyo total de intoxicaciones fuera menor a 80, se definió “Otros” como nivel de referencia. Se permitieron interacciones.

Documentación

El código desarrollado fue comentado para facilitar la lectura, edición y soporte de este (**Anexo 5**), en caso de que personas ajenas a este proyecto deban interactuar con él. El mismo, fue guardado en un repositorio privado de GitHub el cual contiene las versiones que se generaron para conseguir el resultado final.

Resultados

Debido a que este trabajo pretende desarrollar una herramienta para analizar los datos de Intoxicaciones en Costa Rica con estadísticas específicas, los resultados que se muestran a continuación son posibles escenarios que el personal del Centro Nacional de Control de Intoxicaciones de Costa Rica podría considerar, utilizando los datos de intoxicaciones para el 2017 y los datos del censo de 2011.

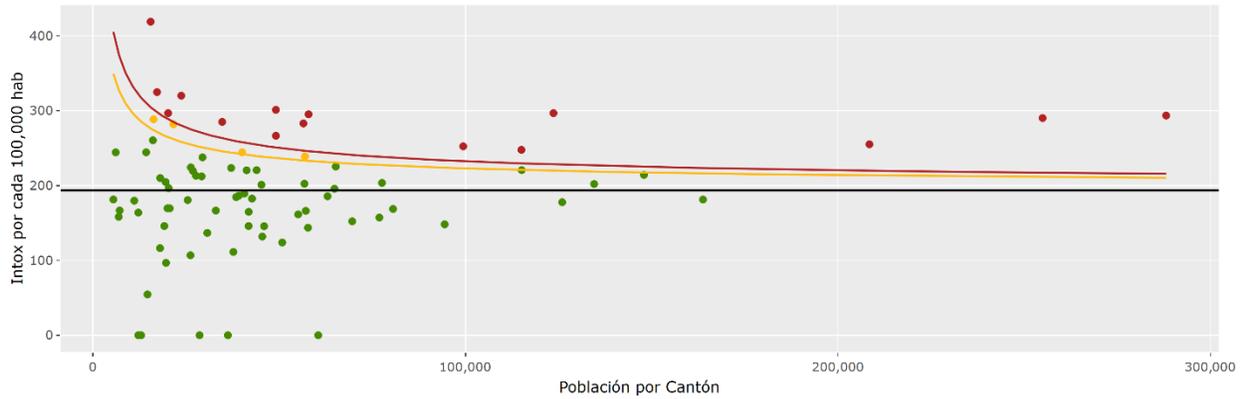
Gráfico de Embudo

A continuación se muestran dos posibles escenarios resultantes del uso de la sección nombrada “Funnel Plot” de la aplicación generada para el análisis de datos de intoxicaciones en Costa Rica.

Escenario I

El gráfico de embudo generado muestra los resultados para todos los cantones de Costa Rica. Se observa 15 cantones: Guatuso, Garabito, Los Chiles, Acosta, Palmares, Santa Ana, Aserrí, Escazú, Montes de Oca, La Unión, Puntarenas, Heredia, Desamparados, Alajuela y San José, de coloración roja dado a que la tasa de intoxicaciones por cada 100,000.00 habitantes en cada caso supera los límites de control del 99% de confianza. Además, se observa 4 cantones: Tarrazú, Belén, Santo Domingo y Moravia, cuyo número de intoxicaciones supera los límites del 95% de confianza mas no los del 99% de confianza. Por otro lado, 63 cantones se encuentran dentro de los límites de control (**Figura 1**).

Figura 1. Gráfico de Embudo: Intoxicaciones en Costa Rica para el 2017 y censo 2011.

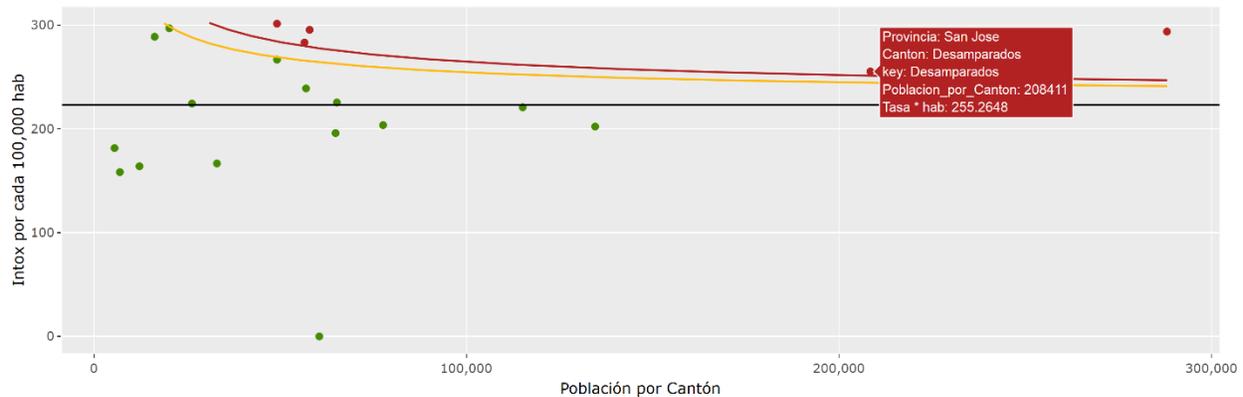


Fuente: Registro de intoxicaciones en Costa Rica, 2017 e INEC, 2011.

Escenario II

El gráfico de embudo generado como segundo ejemplo muestra los resultados para todos los datos de intoxicaciones que fueron reportados al Centro Nacional de Control de Intoxicaciones de Costa Rica en el 2017, limitados de modo de que se observan solo los cantones para la provincia de San José, y su comportamiento de acuerdo a los límites de control. Se observa 5 cantones: Santa Ana, Asserí, Escazú, Desamparados y San José, con una tasa de intoxicaciones por cada 100,000.00 habitantes superior al límite de control del 99%. Específicamente para el caso del Cantón de Desamparados, este supera el límite del 99% de confianza con 255 intoxicaciones por cada 100,000.00 habitantes. No se observan cantones cuyo número de intoxicaciones supere los límites del 95% de confianza mas no los del 99% de confianza, por lo que los restantes 15 cantones se encuentran dentro de los límites de control (**Figura 2**).

Figura 2. Gráfico de Embudo. Intoxicaciones en Costa Rica para el 2017 y censo 2011. Limitado para la provincia de San José.



Fuente: Registro de intoxicaciones en Costa Rica, 2017 e INEC, 2011.

Regresión de Poisson

A continuación se muestran dos posibles escenarios resultantes del uso de la sección nombrada “Regresión de Poisson” de la aplicación generada para el análisis de datos de intoxicaciones en Costa Rica.

Escenario I

La regresión de Poisson para el modelo que incluyó las cuatro provincias del GAM y los reportes de intoxicaciones en personas menores de 19 años, mostró para Alajuela una disminución de las intoxicaciones entre 40.8% y 49.1% con respecto al nivel de referencia (San José). Cartago presentó entre 70.6% y 73.7% menos intoxicaciones que el nivel de referencia (San José) y de manera similar, Heredia presentó entre 68.7% y 74.4% menos intoxicaciones que el nivel de referencia (San José). Lo anterior, con un 95% de confianza para todos los casos (**Cuadro 1**).

Con respecto a los resultados para los grupos etarios, utilizando 0-3 años como el nivel de referencia y un 95% de confianza, se tiene que el rango etario de 4-6 años presentó una disminución de las intoxicaciones entre 69.9% y 74.1%. El rango de 7-12 años presentó una disminución entre 73.7% y 75.8 %. El rango de 13-15 años presentó una disminución entre 82.9% y 87.1 % y el rango de 16-19 años presentó entre 70.3% y 76.4 % menos intoxicaciones que el nivel de referencia. Siendo el rango etario de 13-15 años el que presentó un mayor cambio con respecto al nivel de referencia (**Cuadro 1**).

Cuadro 1. Resultados de la regresión de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.

Variables	IRR	Errores Robustos	LI	LS	Valor p	Significativo (95% de confianza)
(Intercepto)	783.799	15.003	754.937	813.763	0,00	Si
Alajuela	0.549	0.021	0.509	0.592	0,00	Si
Cartago	0.283	0.015	0.256	0.313	0,00	Si
Heredia	0.278	0.008	0.263	0.294	0,00	Si
b. 4-6	0.279	0.011	0.259	0.301	0,00	Si
c. 7-12	0.252	0.005	0.242	0.263	0,00	Si
d.13-15	0.148	0.011	0.129	0.171	0,00	Si
e. 16-19	0.265	0.015	0.236	0.297	0,00	Si

Fuente: Registro de intoxicaciones en Costa Rica, 2017.

La prueba de sobredispersión, utilizando tanto la función lineal como la cuadrática presentó un resultado de dispersión menor a 0, indicando sub-dispersión. Sin embargo, el valor de p es mayor a 0.05, por lo cual no se rechaza la hipótesis nula de que el parámetro de dispersión es igual a cero, utilizando un 95% de confianza (**Cuadro 2**).

Cuadro 2. Resultados de la Prueba de Sobredispersión para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.

Función	Dispersión	Z	Valor p
Func. Lineal	-0.174	-0.591	0.723
Func. Cuadrática	-0.001	-0.984	0.837

Fuente: Registro de intoxicaciones en Costa Rica, 2017.

El resultado de AIC para el modelo de regresión de Poisson presentó un valor de 163.79 (**Figura 3**). Mientras que el resultado del R^2 de McFadden para el modelo de regresión de Poisson presentó un valor de 0.946 (**Figura 4**).

Figura 3. Resultado del AIC para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.



Fuente: Registro de intoxicaciones en Costa Rica, 2017.

Figura 4. Resultado del R^2 de McFadden para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango etario 0-3 años.



Fuente: Registro de intoxicaciones en Costa Rica, 2017.

El resultado del factor de la varianza generalizado para el modelo de Poisson, utilizando 3 grados de libertad para la variable Provincia y 4 grados de libertad para RangoEtario fue de $GFIV^{(1/GL)}$ igual a 1, para ambas variables (**Cuadro 3**).

Cuadro 3. Resultados del factor de inflación de la varianza generalizado para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017 y limitados para las provincias del GAM y la población de 19 años o menos cuyos niveles de referencia son San José y el rango 0-3 años.

Variable	GFIV	GL	$GFIV^{(1/2GL)}$	$GFIV^{(1/GL)}$
Provincia	1	3	1	1
RangoEtario	1	4	1	1

Fuente: Registro de intoxicaciones en Costa Rica, 2017.

Escenario II

Utilizando como nivel de referencia el cantón de San José, se observó que como resultado de la regresión de Poisson se obtuvo que Alajuela presentó en promedio, un 15% menos de intoxicaciones y Heredia un 56.3% menos. En cuanto a la relación entre tóxicos, con respecto al nivel de referencia "Otros", se tiene que se presentaron 97.2% menos intoxicaciones en promedio por el tóxico alacrán que por otros y un 94.5% menos intoxicaciones en promedio, por hipoclorito de sodio (**Cuadro 4**).

Con respecto a los resultados para las interacciones entre los predictores, en promedio las intoxicaciones en Alajuela por Alacranes se presentaron 2.7 veces más que las intoxicaciones en San José por Otros tóxicos. En cuanto a las intoxicaciones en Heredia por Hipoclorito de Sodio también se presentó un aumento con respecto a San José por otros tóxicos, sin embargo un poco menor, siendo este, en promedio, de 1.15 veces más o casi un 15.0% más. Por otro lado, las intoxicaciones por Hipoclorito de

Sodio en Alajuela y en Heredia, mostraron un 28.9% y un 20.1% menos intoxicaciones en promedio que San José por Otros tóxicos, respectivamente (**Cuadro 4**).

Cuadro 4. Resultados de la regresión de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.

Variables	IRR	Errores Robustos	LI	LS	Valor p	Significativo (95% de confianza)
(Intercepto)	781	0,00	781	781	0,00	Si
Alajuela	0.850	0,00	0.850	0.850	0,00	Si
Heredia	0.437	0,00	0.437	0.437	0,00	Si
Alacranes	0.028	0,00	0.028	0.028	0,00	Si
Hipoclorito de Sodio	0.055	0,00	0.055	0.055	0,00	Si
Alajuela:Alacranes	2.673	0,00	2.673	2.673	0,00	Si
Heredia:Alacranes	1.145	0,00	1.145	1.145	0,00	Si
Alajuela:Hipoclorito de Sodio	0.711	0,00	0.711	0.711	0.00	Si
Heredia:Hipoclorito de Sodio	0.799	0,00	0.799	0.799	0,00	Si

Fuente: Registro de intoxicaciones en Costa Rica, 2017.

La prueba de sobredispersión, utilizando tanto la función lineal como la cuadrática presentó un resultado de dispersión menor a 0, indicando tendencia a sub-dispersión. Sin embargo, el valor de p es mayor a 0.05, por lo cual no se rechazó la hipótesis nula de que el parámetro de sobredispersión es igual a cero, utilizando un 95% de confianza (**Cuadro 5**).

Cuadro 5. Resultados de la Prueba de Sobredispersión para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.

Función	Dispersión	Z	Valor p
Func. Lineal	-1	-2198986439808120	1
Func. Cuadrática	-0.002	-2.127	0.983

Fuente: Registro de intoxicaciones en Costa Rica, 2017.

El resultado de AIC para el modelo de regresión de Poisson presentó un valor de 72.71 (**Figura 5**). Por otra parte, el resultado del R² de McFadden para el modelo de regresión de Poisson presentó un valor de 0.98 (**Figura 6**).

Figura 5. Resultado del AIC para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.

72.708
AIC

Fuente: Registro de intoxicaciones en Costa Rica, 2017.

Figura 6. Resultado del R^2 de McFadden para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.

0.983
R² (McFadden)

Fuente: Registro de intoxicaciones en Costa Rica, 2017.

Como resultado del factor de la varianza generalizado para el modelo de Poisson, utilizando 2, 2 y 4 grados de libertad para Cantón, Tóxico y la interacción Cantón:Tóxico, en ese orden, presentó valores de $GFIV^{(1/GL)}$ de 1.1, 2.7 y 1.7 respectivamente (**Cuadro 6**).

Cuadro 6. Resultados del factor de inflación de la varianza generalizado para el modelo de Poisson utilizando las Intoxicaciones en Costa Rica para el 2017, limitados para los cantones San José, Heredia y Alajuela y con los tóxicos agrupados en el nivel Otros si tienen 80 o menos intoxicaciones reportadas.

Variable	GFIV	GL	$GFIV^{(1/2GL)}$	$GFIV^{(1/GL)}$
Canton	1.188	2	1.011	1.090
Toxico	7.301	2	1.644	2.702
Canton:Toxico	8.213	4	1.301	1.693

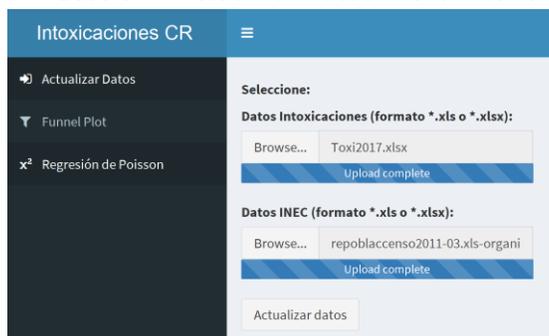
Fuente: Registro de intoxicaciones en Costa Rica, 2017.

Interactividad de la aplicación

Se hace referencia a los resultados de las secciones Gráfico de Embudo y Regresión de Poisson para esta sección.

La primera sección de la interfaz web permitió al usuario ingresar los datos de intoxicaciones del año a analizar, 2017. Así mismo, permitió ingresar los datos del censo 2011. Finalmente, mediante el botón de “Actualizar Datos”, se permitió la carga de los datos y la transición a la siguiente sección (**Figura 7**).

Figura 7. Sección: “Actualizar Datos” de la interfaz web.



La segunda sección de la aplicación, identificada como “funnel plot” permitió, para el **Escenario II** del gráfico de embudo, ajustar los datos de modo que se limitaron y presentaron solamente los cantones correspondientes a la provincia de San José (**Figura 8**).

Figura 8. Filtros en sección: “Funnel Plot” de la interfaz web.



Intoxicaciones CR

Actualizar Datos

Funnel Plot

Resultados por # de habitantes:
100.000

Provincia:
San Jose

Cantón:

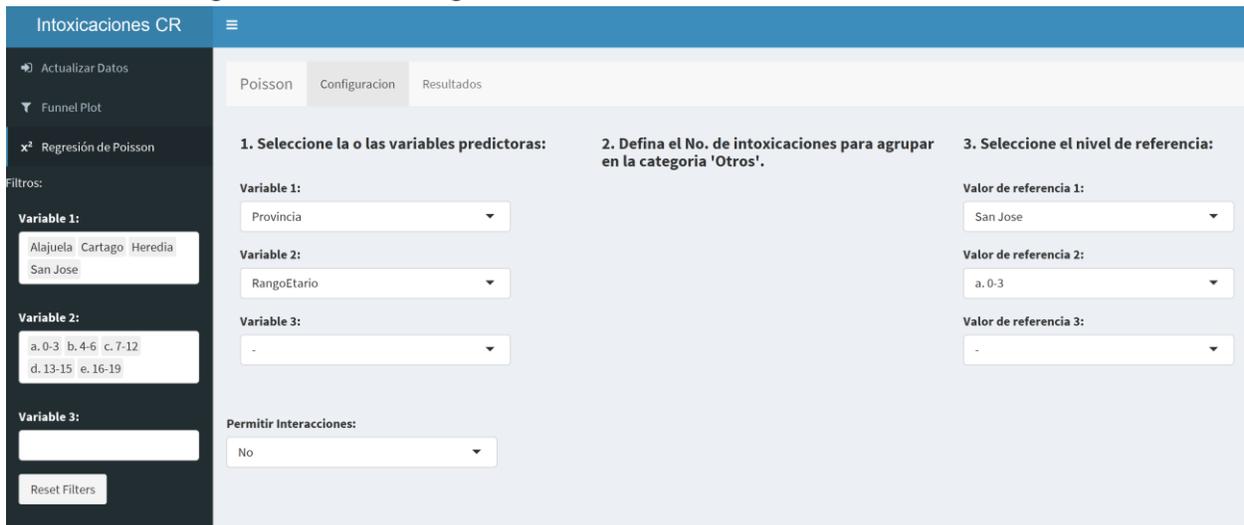
Fecha de Consulta:
17/01/01 a 17/12/31

Edad:
0 99
0 11 22 32 43 54 65 75 86 99

Sexo:
 All
 01.Masculino
 02.Femenino
 03.No anotado

La tercera sección de la aplicación, identificada como “Regresión de Poisson”, permitió definir tanto el modelo para el **Escenario I** (Figura 9), como para el **Escenario II** (Figura 10). Para ambos permitió; la elección de las variables predictoras tras seleccionarlas en el panel central y la definición del nivel de referencia para cada variable. Además permitió la limitación de los datos utilizando el panel de la izquierda (Figura 9 y Figura 10). En el caso del **Escenario II**, se habilitaron interacciones para las variables del modelo, variables de las cuales una permitió la creación de un nivel que incluyó tóxicos con conteos de intoxicaciones iguales o menores a 80 (Figura 10).

Figura 9. Sección: “Regresión de Poisson” de la interfaz web. Escenario I.



Intoxicaciones CR

Actualizar Datos

Funnel Plot

χ² Regresión de Poisson

Filtros:

Variable 1:
Alajuela Cartago Heredia
San Jose

Variable 2:
a. 0-3 b. 4-6 c. 7-12
d. 13-15 e. 16-19

Variable 3:

Reset Filters

Poisson Configuración Resultados

1. Seleccione la o las variables predictoras:

Variable 1:
Provincia

Variable 2:
RangoEtario

Variable 3:
-

2. Defina el No. de intoxicaciones para agrupar en la categoría 'Otros':

Permitir Interacciones:
No

3. Seleccione el nivel de referencia:

Valor de referencia 1:
San Jose

Valor de referencia 2:
a. 0-3

Valor de referencia 3:
-

Figura 10. Sección: “Regresión de Poisson” de la interfaz web. Escenario II.

The screenshot shows the 'Intoxicaciones CR' web interface. The top navigation bar includes 'Actualizar Datos', 'Funnel Plot', and 'Regresión de Poisson'. The main content area is titled 'Poisson' and is divided into three columns for configuration:

- 1. Seleccione la o las variables predictoras:**
 - Variable 1: Canton
 - Variable 2: Toxico1
 - Variable 3: -
- 2. Defina el No. de intoxicaciones para agrupar en la categoría 'Otros':**
 - Categoría 1: 0
 - Categoría 2: 80
- 3. Seleccione el nivel de referencia:**
 - Valor de referencia 1: San Jose
 - Valor de referencia 2: Otros
 - Valor de referencia 3: -

At the bottom, there is a 'Permitir Interacciones:' dropdown menu set to 'Si'.

Discusión de Resultados

Debido a que este trabajo pretende desarrollar una herramienta para analizar los datos de Intoxicaciones en Costa Rica con estadísticas específicas. Los resultados que se discuten a continuación son posibles escenarios que el personal del Centro Nacional de Control de Intoxicaciones de Costa Rica podría considerar, utilizando los datos de intoxicaciones para el 2017 y los datos del censo del 2011.

Gráfico de Embudo

Escenario I

Debido a los resultados mostrados para el gráfico de embudo realizado en el primer escenario, se tiene que para todos los registros del 2017 se presentaron 15 cantones fuera de control en el país (**Figura 1**). Esto se debe a que sus tasas de intoxicación durante el año 2017 encontraron fuera de los límites de confianza del 99% y merecen atención. Por otro lado, hay 4 cantones fueran de los límites de confianza del 95% mas no de los del 99% (**Figura 1**), estos podrían requerir atención, sin embargo, los anteriores son prioritarios (Szklo y Nieto, 2003, Flowers, 2007 y Alan, 2016).

Todos aquellos valores dentro de los límites de confianza (color verde) (**Figura 1**), no requieren atención. Sin embargo, es posible que una vez que se controlen los valores extremos (rojos y amarillos), que los límites de confianza se re-ajusten debido a una nueva media en los datos y que así otros cantones que antes estaban dentro de los límites requieran atención (Appleby, Devlin y Parkin, 2015).

Guatuso ocupó el tercer lugar en intoxicaciones, dado al número de casos por plaguicidas reportados en 2017 (Mayorga, 2018). Montes de Oca, es posible que resalte dado a que en 2017 se presentó un caso de intoxicación masiva en el que resultaron afectadas 15 personas por el consumo de una droga o una mezcla de ellas (Arguedas, 2017).

Para otros cantones, puede que no haya noticias que ejemplifiquen el hecho de que tienen números alarmantes de intoxicaciones con respecto al resto de la población. Sin embargo, hay casos puntuales que son ejemplos de intoxicaciones en el país que ocurrieron, no solo en 2017, pero que como muchas intoxicaciones, pudieron haberse evitado. En 2017 en el cantón de Los Chiles, Alajuela, se reportó el caso de una joven de 15 años que inhaló un pesticida (Jiménez, 2017). En 2017, una niña de dos años

resultó intoxicada por plaguicidas uno de los plaguicidas más tóxicos en el país, llamado Terbufos, el cual tiene una apariencia similar al de una semilla de linaza (Ávalos, 2017).

En 2018, en una escuela de Desamparados se reportó la intoxicación de más de 100 personas entre estudiantes y personal por un virus estomacal que pudo haberse propagado por un mal manejo de los alimentos (Porrás, 2018). En 2018 también se presentaron 7 casos de intoxicaciones en Escazú debido a ingesta de licor vencido (Chaves, 2018). Los anteriores son solo algunos ejemplos de los accidentes que pueden ocasionar que un sector presente un número elevado de intoxicaciones. Sin embargo, debe de considerarse que no solo intoxicaciones accidentales se reportan en nuestro país, también hay intoxicaciones por intentos de suicidio, las cuales representaron un 23.7% de los suicidios en 2016 (Vargas y Espinoza, 2017).

Escenario II

Similar al caso anterior, pero en menor número, se encontraron cantones cuyas tasas de intoxicación superan los límites de confianza, por lo tanto requieren de atención y acciones (Alan, 2016 y Bottle y Aylin, 2016) para eventualmente disminuir la tasa de intoxicaciones. En el caso específico para el cantón de Desamparados, esas 255 intoxicaciones por cada 100,000.00 habitantes, representan un número anormalmente elevado que debe de ser mitigado para que vuelva a estar dentro de los límites de confianza del 95% (**Figura 2**).

El Ministerio de Salud reportó que para el periodo del 2015-2017, las provincias de San José y Puntarenas presentaron un 60.8% de las intoxicaciones en el país, siendo, en 2017, Puntarenas la provincia con mayor cantidad de casos, seguida por San José, Cartago y Alajuela (Mayorga, 2018). El hospital Nacional de Niños, en 2017, indicó que San José y Guanacaste fueron los cantones en los que más intoxicaciones de niños ocurren (Porrás, 2017). Dado a lo anterior es que los resultados de los gráficos de embudo de tanto el **Escenario I** como el **Escenario II**, no sorprenden, ya que indicaron altos números de intoxicaciones en cantones correspondientes a las provincias antes mencionadas.

Para todos los tipos de intoxicaciones se pueden tomar medidas, algunas luego de que la persona ha sido intoxicada con el xenobiótico, tomando en cuenta las características del mismo, la dosis, edad, tiempo de exposición, entre otros (Fernández, 2012). Sin embargo, las que nos resultan más relevantes son aquellas que se enfocan en la prevención de intoxicaciones. Linares y Peñaloza, (2014), indican que el primer paso es tener los datos claros, mismos que en el caso de Costa Rica, se encuentran en el Centro Nacional de Control de Intoxicaciones y se convertirán en información valiosa gracias a esta herramienta. Además Linares y Peñaloza, (2014), recomiendan que debe de realizarse un trabajo en conjunto con diferentes organismos, en el caso de nuestro país podrían ser el CNCI, los distintos centros de salud del país e instituciones educativas, para así poder crear estrategias enfocadas en la prevención.

Como mecanismos de prevención de intoxicaciones se encuentra en primer lugar la educación de la población (Rodríguez et al., 2005 y Fernández, Mancipe y Fernández, 2010), ya que muchas son prevenibles. Rodríguez et al., (2005) mencionan medidas para distintos tipos de intoxicaciones, como las causadas por sobre medicación y las diferencias en la dosificación para niños o adultos, las cuales pueden disminuirse educando. También, se debe estar consciente de las situaciones emocionales de los niños y jóvenes que pueden impulsarlos a cometer actos suicidas (Galicia-Morales, Sánchez-Villegas y Loría-Castellanos, 2016). Además, pese a una buena situación emocional, recomiendan que las sustancias tóxicas no estén al alcance de jóvenes o niños (Rodríguez et al., 2005).

Así mismo, se recomienda en casos de plaguicidas, concientizar a los trabajadores agrícolas sobre el uso y desecho adecuado de los plaguicidas (Fernández et al., 2010 y Chaparro-Narváez y Castañeda-Orjuela, 2015). Medidas nacionales podrían aplicarse, generando políticas para limitar el uso y disponibilidad de plaguicidas (Chaparro-Narváez y Castañeda-Orjuela, 2015). Así mismo, las empresas deberían de promover la responsabilidad ética en la producción industrial y manejo de sustancias tóxicas (Rodríguez et al., 2005). Con respecto a intoxicaciones alimentarias, Mañes y Soriano, (2012), aseguran que la calidad en la preparación y manejo de alimentos puede evitar intoxicaciones. Finalmente, en cuando a animales e insectos, informar sobre el comportamiento y hábitos de los mismos para evitarlos (Vargas y Palacios, 2014).

Regresión de Poisson

Escenario I

Los resultados de Poisson para las provincias del GAM y personas con edades de 19 o menos años presentó, en sus resultados, expresados como tasas de incidencia una disminución del número de intoxicaciones con respecto al nivel de referencia, San José y rango etario de 0-3 años (**Cuadro 1**). Lo anterior se debe a que todos los valores de IRR o de tasa de incidencia son menores a 1 por lo que de acuerdo con Shephard, (2011) y Marcus y Zgierska, (2012) indican disminución de los casos con respecto al nivel de referencia.

Es importante aclarar que los resultados de la tasa de incidencia se expresan distinto a lo presentado en el **Cuadro 1** y que se siguen las recomendaciones de Meyers et al., (2009) quienes para datos con una tasa de incidencia (IRR) o límites menores a 1, utilizan el siguiente cálculo $(1-IRR)*100$ o $(1-LC)*100$, en el caso de los límites de confianza.

La prueba de sobredispersión, indica que no hubo sobredispersión (**Cuadro 2**), ya que para que lo haya se requieren valores de dispersión mayores a 1 junto con un valor de p menor a 0.05 para una confiabilidad del 95% (Hilbe, 2011). En caso de que alguno de los valores hubieran sido mayor a 0 y además tuviera valores de p menor a 0.05, se debería de considerar si el modelo debe de ser ajustado con cambio de variables o filtros (Zuur et al., 2009) o si se deben de considerar otros modelos, como lo es el binomial negativo con su función lineal o cuadrática, según el valor que haya presentado la sobredispersión o el modelo de quasi-Poisson Hayat y Higgins, (2014). En este caso, se podría creer que se presenta el fenómeno de sub-dispersión, sin embargo, se deben de cumplir dos supuestos, que el valor de dispersión sea menor a 0 y que el valor de p sea menor a 0.05, asegurando que hay sub-dispersión con un 95 de confianza (Hilbe, 2014).

El resultado del AIC, 163.79 (**Figura 3**), por sí solo no proporciona gran información (Kramer y Eden, 2016), sin embargo, si se compara este AIC con el de otro modelo ligeramente distinto, como por ejemplo uno similar al utilizado en este escenario pero que incluya más provincias no solo las del GAM o que incluya más rangos etarios, no solo edades de hasta los 19 años, se podrían comparar. Teniendo dos o más posibles modelos, los valores de AIC de estos se podrían contrastar y el de menor AIC se podría considerar el mejor modelo (Westfall y Henning, 2013). Es importante considerar las recomendaciones de Kunert, (2016), quien aconseja no tomar decisiones únicamente con base en el valor del AIC, sino en una serie de estadísticas que indiquen detalles sobre el modelo y su ajuste a los datos.

El valor del R^2 de McFadden, no es un R^2 usual, por tanto no indica si el modelo explica o no la variabilidad de los datos (Kunert, 2016). El R^2 de McFadden se utiliza para comparar el modelo generado

y el modelo nulo y el porcentaje de la variabilidad que explica el modelo con respecto al nulo (Little, 2013). Además, el R^2 de McFadden puede utilizarse para comparar modelos en caso de que pequeños cambios en los datos para el modelo, produzcan cambios en el resultado, en ese caso, se recomendaría el modelo con mayor R^2 de McFadden (Fávero y Belfiore, 2019) ya que diferiría más del modelo que no contiene variables.

Sin embargo, es posible ver valores bajos, lo cual no necesariamente significaría que el modelo es malo o inefectivo, solo que el modelo no se diferencia en mayor proporción del modelo nulo, sino que solo se diferencia en un bajo porcentaje (Little, 2013). El caso para este modelo valor de 0.946 (**Figura 4**) indica que el modelo resultante de **Escenario I** explica un 94.6% más la variabilidad de los datos con respecto al modelo nulo, el cuál de estarse comparando con un modelo con menor R^2 de McFadden, podría ser seleccionado como el mejor (Fávero y Belfiore, 2019). Al igual que para el AIC, es recomendable considerar los otros estadísticos antes de tomar una decisión entre qué modelo es mejor o peor en comparación a otro (Kunert, 2016).

Los resultados del factor de inflación de la varianza fueron generados utilizando la versión generalizada ya que este proyecto utilizó variables con distintos niveles, por lo que resulta importante considerar los grados de libertad dados por los predictores y sus niveles (Fox y Monette, 1992 y Fox, 2002). Valores mayores a 1 para $GFIV^{(1/GL)}$ indican multicolinealidad (Montgomery y Runger, 2018), lo cual quiere decir que hay cierta dependencia entre las variables.

Sin embargo, solo se recomienda eliminar variables y re-evaluar el modelo si el valor de $GFIV^{(1/GL)}$ es mayor a 4 o 5 según Tattar et al., (2016) o mayor a 10 según Montgomery y Runger, (2018). Para los resultados de la regresión de Poisson ejecutada, no se debe de eliminar ninguna de las variables ni re-evaluar el modelo siguiendo las recomendaciones estrictas y no tan estrictas de los autores mencionados previamente, esto porque $GFIV^{(1/GL)}$ fue 1 para ambos predictores (**Cuadro 3**), por lo tanto, se encontraban dentro de los valores aceptables. De ser necesario elegir un criterio, se recomendaría el de Montgomery y Runger, (2018) para definir 10 como el valor de referencia a la hora de analizar los resultados de $GFIV^{(1/GL)}$ por ser más permisivo.

Finalmente, para el **Escenario I** se puede decir que considerando las estadísticas que evalúan el modelo, principalmente del valor del R^2 de McFadden que indica que el modelo explica mucho mejor la variabilidad de los datos que el modelo nulo. El modelo se ajusta bien a los datos y ayuda a visualizar el efecto que tienen las provincias del GAM y los rangos etarios iguales o menores a 19 años en el número de intoxicaciones por año en Costa Rica.

En Bogotá, el hipoclorito de sodio es el producto del hogar que más intoxicaciones presenta en niños de 0-4 años por causas accidentales, seguido por jóvenes de 20-24 años y posteriormente jóvenes de entre los 15-19 años con intenciones suicidas (Linares y Peñaloza, 2014). De otra manera, las intoxicaciones con fines suicidas, reportadas en un centro pediátrico en México, concluyó que la mayoría de casos se presentaban en niños entre 12 y 15 años de edad (Galicia-Morales et al., 2016).

Otros estudios indican que la mayoría de las intoxicaciones accidentales ocurren en niños menores de 5 años, de los que predominan los de sexo masculino (Prado, 2011, Oliva y Leiva, 2012 y Abente, Blanes, y de Núñez, 2013). La causa principal varía según estudio, en algunos, resaltan los medicamentos (Prado, 2011), gases tóxicos (Rodríguez et al., 2005, Abente, Blanes, y de Núñez, 2013) y plaguicidas (Oliva y Leiva, 2012). Opuesto al comportamiento anterior, un hospital en Ecuador reportó

para el 2016 que la mayoría de casos de intoxicaciones se presentaban por ingesta accidental de parte de hombres entre los 20-49 años (Segura et al., 2016).

Escenario II

Le regresión de Poisson mostrada en el **Escenario II** ejemplifica el uso de interacciones lo cual significa que además de las variables seleccionadas, se crean combinaciones de los niveles de las variables, para determinar el impacto que tienen en pares (Millar, 2011). En todos los casos, menos en Alajuela:Alacranes y en Heredia:Alacranes se mostraron disminuciones en el número de casos de intoxicaciones de acuerdo al nivel de referencia (**Cuadro 4**). Lo anterior se respaldó gracias Shephard, (2011) y Marcus y Zgierska, (2012) los cuales indican que valores la tasa de incidencia o IRR menores a 1 indican una disminución de los casos y valores mayores a 1 indican un aumento de los casos por año.

Dado a lo anterior es que también se puede comprender que los valores para las interacciones Alajuela:Alacranes y Heredia:Alacranes, indiquen que en estas dos provincias se presentan más casos de intoxicaciones por alacranes con respecto al nivel de referencia, San José, considerando otros tóxicos (**Cuadro 4**).

Se debe de tomar en cuenta que los resultados mencionados, tanto en este escenario como el anterior, no son relativos a la población del cantón o provincia de la que se habla. Es decir, son los valores con respecto al total de las intoxicaciones reportadas, no con respecto a las intoxicaciones de cada cantón o provincia. Un ejemplo de lo anterior sería un cantón A con 10 intoxicaciones por cada 100 habitantes y un cantón B con 10 intoxicaciones por cada 50 habitantes. La regresión de Poisson evaluada indicaría para ambos casos un mismo IRR, ya que en ambos el número absoluto de intoxicaciones es igual, 10. Sin embargo, al observar la población en cada cantón, el caso para el cantón B es más severo, ya que hay dos veces más intoxicaciones que en el caso A, lo cual sería evidente si se considerara la frecuencia relativa. Lo anterior podría darse utilizando un contrapeso u *offset* en la regresión de Poisson tal y como lo ejemplifican Dunn y Smyth, (2018).

Además, se debe aclarar que los resultados de la tasa de incidencia se expresan distinto a lo presentado en el **Cuadro 4** ya que su interpretación se basa en las recomendaciones de Meyers et al., (2009) y Harris et al., (2012) quienes para datos con una tasa de incidencia (IRR) o límites menores a 1 utilizan la siguiente fórmula $(1-IRR)*100$, obteniendo un valor porcentual. Para valores de IRR mayores a 1, se expresan como cantidad de veces más o $(IRR-1)*100$ en caso de preferir un porcentaje de cambio.

En cuanto a los valores de la prueba de sobredispersión (**Cuadro 5**), el criterio de información de Akaike (**Figura 5**), el R^2 de McFadden (**Figura 6**) y el factor de inflación de la varianza generalizado (**Cuadro 6**), se obtuvo comportamientos similares a los obtenidos anteriormente para el **Escenario I**. Las explicaciones dadas anteriormente aplican para estos casos por lo cual se omiten con tal de no repetir información innecesariamente.

Dado a lo anterior, se tiene que para el **Escenario II**, de igual manera que para el **Escenario I**, tomando en cuenta las estadísticas que evalúan el modelo, que el modelo de Poisson ejecutado explica la variabilidad de los datos con respecto al modelo nulo, por lo tanto, es diferente al modelo nulo y nos ayuda a visualizar el efecto que tienen los Cantones, (San José, Heredia y Alajuela), los tóxicos (con conteos menores a 80 intoxicaciones (Otros), el Hipoclorito de Sodio y los Alacranes) y las interacciones entre ellos en el número de intoxicaciones para el 2017 en Costa Rica.

Curiosamente, en México las picaduras de alacranes son la décima causa de intoxicaciones de niños y adolescentes de 9 a 14 años y es considerado, en algunos estados, como Durango y Morelos, como un problema de salud pública (Vargas y Palacios, 2014 y Santibáñez-López et al., 2016). Vargas y Palacios, (2014), indican que la población más vulnerables son los niños ya que son picados en las noches. Además, otorgan una serie de recomendaciones para evitar estas picaduras, por lo que podrían considerarse, para el caso de Alajuela y Heredia, sin importar la edad de las personas, ya que las mismas aplican para toda la población. Por otro lado, podría generarse otro modelo de Poisson y esta vez, se podría agregar RangoEtario como una variable más y determinar si hay uno o varios grupos con mayor vulnerabilidad.

El hipoclorito de sodio, es encontrado en los hogares como blanqueadores, purificadores de agua y productos de limpieza, popularmente conocido como cloro (Peña y Zuluaga, 2017). Las intoxicaciones con este pueden ocurrir por ingesta o por inhalación de los vapores al ser mezclados con amoníaco o ácido muriático (Saracco, 2016). Dada a la facilidad con la que se consigue el hipoclorito de sodio es que se presenta gran cantidad de intoxicaciones, ya que a veces por deficientes métodos de almacenamiento son ingeridos por niños. En un estudio realizado en 2014 en Bogotá se obtuvo que este encabeza la lista de intoxicaciones por productos del hogar y que la mayoría de estas intoxicaciones se dan en niños de 0 – 4 años (Linares y Peñaloza, 2014).

De acuerdo a los resultados obtenidos, se podrían generar campañas educativas, que creen un impacto en la sociedad. Por ejemplo, si un área de nuestro país presenta un alto número de intoxicaciones por alacranes, podría generarse campañas de educación en las que se informe a la población sobre los hábitos del alacrán y medidas preventivas como no andar descalzos, sacudir ropa y zapatos antes de utilizarlos, entre otros (Vargas y Palacios, 2014). En el caso del hipoclorito de sodio, se pueden crear campañas en donde se informe a grandes y niños del problema que estos presentan. Además, que deben de ser almacenados fuera del alcance de los niños, que los recipientes deben de ser desechados adecuadamente y que nunca se deben de utilizar recipientes de alimentos para almacenarlos, entre otras medidas (Linares y Peñaloza, 2014).

Interactividad de la aplicación

La interactividad de la aplicación fue utilizada para permitir la visualización de los gráficos y cuadros que se presentan en este trabajo. Todo gracias al paquete *shiny* 1.1.0 ya que este integra código de HTML y JavaScript (Chang et al., 2019). HTML (HyperText Markup Language) significa lenguaje de marcas de hipertexto y permite que un usuario pueda interactuar con lo que se ve en la pantalla al indicarle a un navegador web cómo desplegar una página (Willard, 2013). JavaScript, es un lenguaje de programación creado para complementar páginas web otorgándoles interactividad (Azaustre, 2016).

Mediante el uso de *shiny* y considerando sus fortalezas al integrar lenguajes para la creación de aplicaciones (Beeley, 2013) y el poder para el análisis estadístico de R (Gardener, 2012 y Hector, 2017) es que se permitió la creación de la aplicación en la que dinámicamente se pueden realizar gráficos de embudos y modelos de Poisson de acuerdo a los datos de interés, en muy pocos segundos, mediante la interactividad de la aplicación (Resnizky, 2015). Así mismo, fortalezas propias de *shiny* y paquetes que con este se relacionan, permitieron la creación de mayor interactividad para así mejorar la experiencia del usuario (Chang y Borges, 2018, Sievert et al., 2019 y Xie, Cheng y Tan, 2019).

Es por lo anterior que los resultados de tanto la sección del gráfico de embudo como la de la regresión de Poisson muestran datos diferentes. La interactividad de la aplicación otorga la posibilidad

de, con pequeños o grandes cambios, obtener diferentes escenarios, analizar distintas situaciones y tomar decisiones (**Figuras 7-10**).

Tomando como ejemplos los escenarios evaluados en este trabajo, se logró interactividad al habilitar las diferentes secciones para separar la carga o actualización de datos (**Anexo 1**), el gráfico de embudo el cual varió de escenario a escenario por una simple selección de filtros (**Anexo 2**) y la regresión de Poisson, la cual mostró dos escenarios completamente distintos al utilizar variables predictoras distintas, filtros y valores de referencia variables de acuerdo al número de intoxicaciones (en el caso del **Escenario II**) (**Anexo 3 y Anexo 4**). Además para facilitar la comprensión de los resultados del modelo de Poisson por parte del usuario se agregaron mensajes con guías para posibles interpretaciones, las cuales se basan en la literatura utilizada en este trabajo (**Anexo 4**).

Conclusiones y Recomendaciones

Se creó un gráfico de embudo que permite a los usuarios identificar tasas anormales de intoxicaciones para los distintos cantones de Costa Rica.

Se generó un modelo de Poisson en el que los usuarios pueden identificar variables que están influyendo en el número de intoxicaciones que se presentan en Costa Rica cada año.

Se desarrolló una interfaz web en la que se permite al usuario interactuar y así poder modificar los datos que se utilizan tanto para el gráfico de embudo como para el modelo de Poisson. Además se proporcionó la funcionalidad para modificar las variables predictoras, los niveles a ser incluidos y otras configuraciones relacionadas con los niveles de referencia.

Se recomienda valorar cuidadosamente los resultados de la regresión de Poisson y en caso de existir sobredispersión, ajustar el modelo, de no ser posible, ignorar los resultados y valorar opciones como el modelo de quasi-Poisson o el modelo binomial negativo y luego comparaciones entre ambos para determinar que se utilice el modelo óptimo. Similar en caso de sobredispersión, utilizando el modelo de Hurdle como alternativa.

Bibliografía

Abente, M.L., Blanes, M. y de Núñez, S. (2013). Perfil epidemiológico de intoxicaciones pediátricas registradas en el Centro Nacional de Toxicología (MSP y BS). *Dialnet*, 40(1), 29-34.

Alan, J. (2016). *Statsnotes: Some Statistics for Management Problems*. Londres, Inglaterra. World Scientific.

Alves, R., de Morais, O.L., Rovey, M., Cortez-Escalante, J.J., de Pinho, T.A., Weirich, C.F., Mangueira, M., Ferreira, J., Sobrinho, G., de Jesus, L., Grego, L. (2019). Epidemiology of Self-Reported Diabetes Mellitus in the State of Maranhão, Northeastern Brazil: Results of the National Health Survey, 2013. *Int. J. Environ. Res. Public Health*, 16(1), <https://doi.org/10.3390/ijerph16010047>

Appleby, J., Devlin, N., Parkin, D. (2015). *Using Patient Reported Outcomes to Improve Health Care*. Susse del Oeste, Inglaterra. John Wiley & Sons.

Arguedas, C. (2017). Intoxicación masiva con droga en Montes de Oca queda sin Castigo. *La Nación*. Recuperado el 1 de Mayo de 2019, desde: <https://www.nacion.com/sucesos/narcotrafico/intoxicacion-masiva-con-droga-en-montes-de-oca-queda-sin-castigo/IJPHLKCBJBGLRKUOEQYZQ67N2A/story/>

Arroyo, Y., Sala, A.E., y Arias, F. (2014). Tendencia de la mortalidad por casos de intoxicaciones en Costa Rica. *Revista Costarricense de Salud Pública*, 23(1), 58–62.

Ávalos, A. (2017). Inusual intoxicación con plaguicida paraliza a niña de dos años. *La Nación*. Recuperado el 1 de Mayo de 2019, desde: <https://www.nacion.com/el-pais/salud/inusual-intoxicacion-con-plaguicida-paraliza-a-nina-de-dos-anos/7GK4KQ6KHJHRBCZWN5MYELHHCUCU/story/>

Azaustre, C. (2016). *Aprendiendo JavaScript*. Madrid, España. Carlosazaustre.es

Barbosa A.M., Brown J.A., Jiménez-Valverde A., Real R. (2016). *modEvA: Model Evaluation and Analysis*. Paquete de R versión 1.3.2. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/modEvA/modEvA.pdf>

Beeley, C. (2013). *Web Application Development with R using Shiny*. Birmingham, Reino Unido. Packt Publishing.

Bergman, J.E., Loane, M., Vrijheid, M., Pierini, A., Nijman, R.J., Addor, M., Barisic, I., Béres, J., Braz, P., Budd, J., Delaney, V., Gatt, M., Khoshnood, B., Klungsoyr, K., Martos, C., Mullaney, C., Nelen, V., Neville, A.J., O'Mahony, M., Queisser-Luft, A., Randrianaivo, H., Rissmann, A., Rounding, C., Tucker, D., Wellesley, D., Zymak-Zakutnia, N., Bakker, M.K. y de Walle, H.E. (2015). Epidemiology of hypospadias in Europe: a registry-based study. *World Journal of Urology*, 33(12), 2159-2167.

Bottle, A. y Aylin, P. (2016). *Statistical Methods for Healthcare Performance Monitoring*. Florida, Estados Unidos. CRC Press.

Cabo, J. (2014). *Gestión de la calidad en las organizaciones sanitarias*. Madrid, España. Ediciones Díaz de Santos.

Cameron, A.C. y Trivedi, P.K. (2009). *Microeconomics Using Stata*. Texas, Estados Unidos. Stata Press.

Capel, J. y Yuste, F.J. (2016). *Manual de Prácticas de Bioinformática*. Almería, España. Editorial Universidad de Almería.

- Chakraborty, S., Komatsu, K., Roberts, M., Collins, J., Beggs, J., Turabelidze, G., Safranek, T., Maillard, J., Bell, L.J., Young, D., Marsden-Haug, N., Klos, R.F. y Dworkin, M.S. (2015). The Descriptive Epidemiology of Yersiniosis: A Multistate Study, 2005–2011. *Public Health Reports*, 130(3), <https://doi.org/10.1177/003335491513000314>
- Chambers, J.M. (2017). *Extending R*. Florida, Estados Unidos. CRC Press.
- Chang, W. y Borges, B. (2018). shinydashboard: Create Dashboards with “Shiny”. Paquete de R versión 0.7.1. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/shinydashboard/shinydashboard.pdf>
- Chang, W., Cheng, J., Allaire, J.J., Xie, Y. y McPherson, J. (2019). shiny: Web Application Framework for R. Paquete de R versión 1.3.2. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/shiny/shiny.pdf>
- Chaparro-Narváez, P. y Castañeda-Orjuela, C. (2015). Mortalidad debido a intoxicación por plaguicidas en Colombia entre 1998 y 2011. *Biomédica*, 35(1), 90-102.
- Charpentier, A. (2014). *Computational Actuarial Science with R*. Montreal, Canada. CRC Press.
- Chaves, K. (2018). OIJ: Siete muchachos intoxicados tras ingerir licor vencido en fiesta de cumpleaños. *La Nación*. Recuperado el 1 de Mayo de 2019, desde: <https://www.nacion.com/sucesos/accidentes/supuesta-intoxicacion-con-alcohol-deja-seis/6GQLSJZVOZHXFHRDGETWEHCS3M/story/>
- Chekanov, S.V. (2016). *Numeric Computation and Statistical Data Analysis on the Java Platform*. Illinois, Estados Unidos. Springer.
- Cordeiro, R.L.F., Faloutsos, C. y Traina, C. (2013). *Data Mining in Large Sets of Complex Data*. Londres, Inglaterra. Springer Science & Business Media.
- Correa, J.C. y Barrera, C.J. (2018). *Introducción a la Estadística Bayesiana*. Medellín, Colombia. Instituto Tecnológico Metropolitano.
- Cotton, R. (2013). *Learning R: A Step-by-Step Function Guide to Data Analysis*. California, Estados Unidos. O’Reilly Media.
- da Cruz, P., Toporcov, T.N., Rotundo, L.D., Biazevic, M.G., Brasileiro, R.S., de Carvalho, M.B., Kowalski, L.P., Antunes, J.L. (2012). Food restrictions of patients who are undergoing treatment for oral and oropharyngeal cancer. *European Journal of Oncology Nursing*, 16(3), 253-257.
- de la Torre, M. (2014). Intoxicaciones más Frecuentes. *Pediatría Integral*, 18(5), 280-290.
- Dunn, P.K. y Smyth, G.K. (2018). *Generalized Linear Models With Examples in R*. Nueva York, Estados Unidos. Springer.
- Eayres, D. (2010). *Technical Briefing 3: Commonly used public health statistics and their confidence intervals*. North Yorkshire, Inglaterra. Association of Public Health Observatories.
- Fávero, L.P. y Belfiore, P. (2019). *Data Science for Business and Decision Making*. Londres, Inglaterra. Academic Press.

- Fernández, D.G., Mancipe, L.C. y Fernández, D.C. (2010). Intoxicación por organofosforados. *Revista Med*, 18(1), 84-92.
- Fernández, R. (2012). *Manual de prevención de riesgos laborales para no iniciados (2da Ed.)*. Alicante, España. Editorial Club Universitario.
- Flowers, J. (2007). *Technical Briefing 2: Statistical process control methods in public health intelligence*. North Yorkshire, Inglaterra. Association of Public Health Observatories.
- Fox, G., Negrete-Yankelevich, S. y Sosa, V. (2015). *Ecological Statistics*. Oxford, Inglaterra. Oxford University Press.
- Fox, J. (2002). *An R and S-Plus Companion to Applied Regression*. California, Estados Unidos. SAGE.
- Fox, J. y Monette, G. 1992. Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, 87(417), 178-183.
- Fox, J., Wesiberg, S. y Price, B. (2018). *car: Companion to Applied Regression*. Paquete de R version 3.0-2. Recuperado el 1 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/car/car.pdf>
- Friendly, M. y Meyer, B. (2016). *Discrete Data Analysis with R: Visualization and Modeling Techniques for Categorical and Count Data*. Florida, Estados Unidos. CRC Press.
- Galicia-Morales, Y., Sánchez-Villegas, M. y Loría-Castellanos, J. (2016). Intoxicaciones con fines suicidas atendidas en el servicio de urgencias de pediatría del Centro Médico la Raza. *Sociedad Mexicana de Medicina de Emergencia*, 8(3), 81-89.
- Ganguly, K. (2017). *R Data Analysis Cookbook (2da Ed.)*. Birmingham, Inglaterra. Packt Publishing Ltd.
- Gardener, M. (2012). *Beginning R: The Statistical Programming Language*. Indianápolis, Estados Unidos. John Wiley & Sons.
- Gentile, N., Khanna, S., Loftus, E., Smyrk, T., Tremaine, W., Harmsen, S., Zinsmeister, A., Kammer, P., Pardi, D. (2014). The Epidemiology of Microscopic Colitis in Olmsted County from 2002 to 2010: A Population-Based Study. *Clinical Gastroenterology and Hepatology*, 12(5), 838-842.
- Ghiggi, M. (2014). *Manual de primeros auxilios infantiles: Guías para padres y cuidadores*. T-edito.com.
- Girón, L.E. (2017). *Econometría aplicada usando stata 13*. Bogotá, Colombia. Pontificia Universidad Javeriana.
- Guitart, R. (2014). *Tóxicos: Los enemigos de la vida*. Barcelona, España. Ediciones UAB.
- Guzmán-Plazola, P., Guevara-Gutiérrez, R.D., Olguín-López, J.L. y Mancilla-Villa, O.R. (2016). Perspectiva Campesina, intoxicaciones por plaguicidas y uso de agroquímicos. *Idesia*, 34(3), 67-78, <http://dx.doi.org/10.4067/S0718-34292016000300009>
- Harris, T., Yang, Z. y Hardin, J.W. (2012). Modeling underdispersed count data with generalized Poisson regression. *The Stata Journal*, 12(4), 736-747.
- Hawkins, D.M. y Olwell, D.H. (2012). *Cumulative Sum Charts and Charting for Quality Improvement*. Nueva York, Estados Unidos. Springer.

- Hayat, M. J., & Higgins, M. (2014). Understanding Poisson Regression. *Journal of Nursing Education*, 53(4), 207–215. doi:10.3928/01484834-20140325-04
- Hector, A. (2017). *The New Statistics with R: An Introduction for Biologists*. Oxford, Reino Unido. Oxford University Press.
- Hernández, G., Moreno, A., Zaragoza, F. y Porras, A. (2010). *Tratado de Medicina Farmacéutica*. Madrid, España. Medica Panamericana.
- Hilbe, J.M. (2011). *Negative Binomial Regression (2da Ed.)*. Nueva York, Estados Unidos. Cambridge University Press.
- Hilbe, J.M. (2014). *Modeling Count Data*. Nueva York, Estados Unidos. Cambridge University Press.
- Ieva, F. y Paganoni, A.M. (2014). Detecting and visualizing outliers in provider profiling via funnel plots and mixed effect models. *Health Care Management Science*, 18(2), 166–172.
- Instituto Nacional de Estadística y Censos (INEC). 2011. Recuperado el 17 de Mayo de 2018, desde: <http://www.inec.go.cr/poblacion/estimaciones-y-proyecciones-de-poblacion>
- Jiménez, S. (2017). Joven de 15 años inhala sustancia tóxica en Los Chiles. *La Prensa Libre*. Recuperado el 1 de Mayo de 2019, desde: <https://www.laprensalibre.cr/Noticias/detalle/127614/joven-de-15-anos-inhala-sustancia-toxica-en-los-chiles>
- Joyanes, L. (2016). *Big Data, Análisis de grandes volúmenes de datos en organizaciones*. México D.F., México. Alfaomega Grupo Editor.
- Kaminski, R.J. (2004). *The Murder of Police Officers*. Michigan, Estados Unidos. LFB Scholarly Pub.
- Kleiber, C. y Zeileis, A. (2018). *AER: Applied Econometrics with R*. Paquete de R versión 1.2-6. Recuperado el 1 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/AER/AER.pdf>
- Kleinman, K. y Horton, N.J. (2014). *SAS and R: Data Management, Statistical Analysis and Graphics (2da Ed.)*. Massachusetts, Estados Unidos. CRC Press.
- Kramer, M.A. y Eden, U.T. (2016). *Case Studies in Neural Data Analysis*. Londres, Inglaterra. MIT Press.
- Kurnet, J. (2016). *Members of the European Parliament on the Web*. Baden-Baden, Alemania. Nomos.
- Libby, T. (Ed.) y Thorne, L. (Ed.). (2017). *The Routledge Companion to Behavioural Accounting Research*. Routledge.
- Linares, J.A. y Peñaloza, J.A. (2014). Caracterización de intoxicaciones con productos de higiene de uso doméstico en Bogotá D.D. reportadas al SIVIGILA entre los años 2008 y 2012 (tesis de grado). Universidad de Ciencias Aplicadas y Ambientales.
- Lim, A. y Tjhi, W. (2015). *R High Performance Programming*. Birmingham. Reino Unido. Packt Publishing.
- Little, T.D. (2013). *The Oxford Handbook of Quantitative Methods (2da Ed.)*. Oxford, Estados Unidos. Oxford University Press.

Mañes, J. y Soriano, J.M. (2012). La calidad como prevención de las intoxicaciones alimentarias: Toxicología alimentaria. Madrid, España. Ediciones Díaz de Santos.

Marcus, M. y Zgierska, A. (2012). Mindfulness-Related Treatments and Addiction Recovery. Nueva York, Estados Unidos. Routledge.

Martínez-Valenzuela, C. y Gómez-Arroyo, S. (2007). Riesgo genotóxico por exposición a plaguicidas en trabajadores agrícolas. *Revista internacional de contaminación ambiental*, 23(4), 185-200.

Mayorga, K. (2018). Memoria Institucional 2014-2018. Ministerio de Salud de Costa Rica. Recuperado el 1 de Mayo de 2019, desde: https://www.ministeriodesalud.go.cr/sobre_ministerio/memorias/memoria_2014_2018/dpi_memoria%20_ministerio_salud_2014_2018_1.pdf

Meyers, D.G., Neuberger, J.S. y He, J. (2009). Cardiovascular Effect of Bans on Smoking in Public Places. *Journal of the American College of Cardiology*, 54(14), 1249-1255.

Millar, R.B. (2011). Maximun Likelihood Estimation and Inference: with Examples in R, SAS and ADMB. Sussex Occidental, Reino Unido. John Wiley & Sons Ltd.

Mintegi, S., Esparza, M.J., González, J.C., Rubio, B., Sánchez, F., Vila, J.J., Yagüe, F. y Benítez, M.T. (2015). Recomendaciones sobre la prevención de intoxicaciones. *Anales de Pediatría*, 83(6), 440.e1-440.e5, <https://doi.org/10.1016/j.anpedi.2015.01.003>

Montgomery, D.C. y Runger, G.C. (2018). Applied Statistics and Probability for Engineers (7ma Ed.). Arizona, Estados Unidos. John Wiley & Sons.

Morreale, P. (Ed.), Goncalves, A. (Ed.) y Silva, C. (Ed.). (2015). Analysis and Visualization of Large-Scale Time Series Network Data. Nueva Jersey, Estados Unidos. Edu.

Morton, Q., Mengersen, K., Whitby, M. y Playford, G. (2013). Statistical Methods for Hospital Monitoring with R. Sussex Occidental, Reino Unido. John Wiley & Sons Ltd.

Oliva, A.D. y Leiva, A.L. (2012). Comportamiento de los pacientes con intoxicaciones agudas de la provincia de Villa Clara. *Revista Cubana de Toxicología*, 1(1).

PennState. (2018). Poisson Regression Model. Pensilvania: The Pennsylvania State University. Recuperado el 12 de Mayo de 2018, desde: <https://newonlinecourses.science.psu.edu/stat504/node/168/>

Peña, L.M. (Ed.) y Zuluaga, A.F. (Ed.). (2017). Protocolos de manejo del paciente intoxicado (2da Ed.). Antioquia, Colombia. Universidad de Antioquia.

Porras, K. (2017). Cada 3 horas se intoxicó un niño en 2016. Monumental. Recuperado el 1 de Mayo de 2019, desde: <http://www.monumental.co.cr/2017/06/28/cada-3-horas-se-intoxico-un-nino-en-el-2016/>

Porras, K. (2018). 105 estudiantes no regresaron a escuela de Desamparados por intoxicación de semana pasada. Monumental. Recuperado el 1 de Mayo de 2019, desde: <http://www.monumental.co.cr/2018/02/26/105-estudiantes-no-regresaron-escuela-de-desamparados-por-intoxicacion-de-semana-pasada/>

- Prado, Y., Vizcaíno, M., Abeledo, C., Prado, E. y Leiva, O. (2011). Intoxicaciones agudas en pediatría. *Revista Cubana de Pediatría*, 83(4), 356-364.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rabe-Hesketh, S. y Skrondal, A. (2008). *Multilevel and Longitudinal Modeling Using Stata* (2da Ed.). Texas, Estados Unidos. Stata Press.
- Registro de intoxicaciones en Costa Rica (Filtrado). (2017). Toxi2017. Recuperado el 25 de enero del 2019, desde: Centro Nacional de Control de Intoxicaciones de Costa Rica, San José, Costa Rica.
- Resnizky, H.G. (2015). *Learning Shiny*. Birmingham, Reino Unido. Packt Publishing.
- Rodríguez, G. (2007). *Generalized Linear Models: Poisson Models for Count Data*. Nueva Jersey, Estados Unidos. Princeton University.
- Rodríguez, L., Wilkins, A., Olvera, R. y Silva, R. (2005). Panorama epidemiológico de las intoxicaciones en México, *Medicina Interna de México*, 21(2), 123-132.
- Ryan, T.P. (2013). *Sample Size Determination and Power*. New Jersey, Estados Unidos. John Wiley & Sons.
- Salinas-Rodríguez, A., Manrique-Espinoza, B. y Sosa-Rubí, S.G. (2009). Análisis estadístico para datos de conteo: aplicaciones para el uso de servicios de salud. *Salud Pública de México*, 51(5), 397-406.
- Santibáñez-López, C.E., Francke, O.F., Ureta, C y Possani, L.D. (2016). Scorpions from Mexico: from Species Diversity to Venom Complexity. *Toxins*, 8(1), <https://doi.org/10.3390/toxins8010002>
- Saracco, S. (2016). *Recomendaciones para la Atención d las Intoxicaciones por Lavandina/Cloro*. Mendoza Argentina. Centro de Información y Asesoramiento Toxicológico.
- Schmuller, J. (2016). *Statistical Analysis with R*. Nueva Jersey, Estados Unidos. John Wiley & Sons.
- Segura, M., Lam, A., Santos, J., López, M. y Sanmartin, D. (2016). Incidencia de Intoxicaciones: Un caso en Hospital de Ecuador. *Ciencia Unemi*, 9(19), 77-83.
- Shephard, R.J. (2011). *Year Book of Sports Medicine 2011*. Pennsylvania, Estados Unidos. Elsevier.
- Sievert, C., Parmer, C., Hocking, T., Chamberlain, S., Ram, K., Corvellec, M. y Despouy, P. (2019). plotly: Create Interactive Web Graphics via "plotly.js". Paquete de R versión 4.9.0. Recuperado el 1 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/plotly/plotly.pdf>
- Sribney, W. y Wiggins, V. (2019). Standard errors, confidence intervals, and significance tests for ORs, HRs, IRRs, and RRRs. *Stata*. Recuperado el 1 de Mayo de 2019, desde: <https://www.stata.com/support/faqs/statistics/delta-rule/>.
- Staggs, V.S. (2013). Nurse staffing, RN mix, and assault rates on psychiatric units. *Research in Nursing and Health*, 36(1), 26-37.
- Sullivan, L.M. (2011). *Essentials of Biostatistics in Public Health* (2da Ed.). Massachusetts, Estados Unidos. Jones & Bartlett Publishers.

- Szklo, M. y Nieto, J. (2003). *Epidemiología intermedia: conceptos y aplicaciones*. Madrid, España. Ediciones Díaz de Santos.
- Tattar, P.N., Ramaiah, S. y Manjunath, B.G. (2016). *A Course in Statistics with R*. Chennai, India. John Wiley & Sons.
- Teutonico, D. (2013). *Instant R Starter*. Birmingham. Reino Unido. Packt Publishing.
- Theisen, S., Drabik, A. y Stock, S. (2012). Pressure ulcers in older hospitalized patients and its impact on length of stay: a retrospective observational study. *Journal of Clinical Nursing*, 21(3-4), 380-387.
- Unwin, A. (2016). *Graphical Data Analysis with R*. Florida, Estados Unidos. CRC Press.
- Vargas, A. y Palacios, V.P. (2014). *Educación para la Salud*. Distrito Federal, México. Grupo Editorial Patria.
- Vargas, M.P. (2016). Intoxicación por acetaminofén en adultos. *Medicina Legal de Costa Rica*, 33(1), 103-109.
- Vargas, M. y Espinoza, K. (2017). Suicidios En Costa Rica Durante Los Años 2015 Y 2016. *Medicina Legal de Costa Rica*, 34(1). Recuperado el 11 de Mayo de 2019, desde: <https://www.scielo.sa.cr/pdf/mlcr/v34n1/1409-0015-mlcr-34-01-00035.pdf>
- Vargas, Y. (2015). Exposición a agroquímicos y creencias asociadas a su uso en la cuenca hidrográfica del Río Morote, Guanacaste, Costa Rica: Un estudio de casos. *Ciencia & Trabajo*, 17(52), 54-68, <http://dx.doi.org/10.4067/S0718-24492015000100011>
- Velázquez-Perea, L.E. (2017). *Estadística Descriptiva y Probabilidad con Excel*. México. CreativeCommons.
- Viales, G. (2014). Intoxicaciones por Paraquat. *Medicina Legal de Costa Rica*, 31(2), 88-94.
- Vilar, J.F. (2005). *Control estadístico de los procesos (SPC)*. Madrid, España. FC Editorial.
- Vo, P., Czygan, M., Kumar, A. and Raman, K. (2017). *Python: Data Analytics and Visualization*. Birmingham, Reino Unido. Packt Publishing.
- Ward, H., Toledano, M.B., Shaddick, G., Davies, B. and Elliot, P. (2012). *Oxford Handbook of Epidemiology for Clinicians*. Oxford, Reino Unido. OUP Oxford.
- Watanitanon, A., Lyons, V.H., Lele, A.V., Krishnamoorthy, V., Chaikittisilpa, N., Chandee, T., Vavilala, M.S. (2018). Clinical Epidemiology of Adults With Moderate Traumatic Brain Injury. *Critical Care Medicine*, 46(5), 781-787, <https://doi.org/10.1097/CCM.0000000000002991>
- Weaver, C.G., Ravani, P., Oliver, M.J., Austin, P.C. y Quinn, R.R. (2015). Analyzing hospitalization data: potential limitations of Poisson regression. *Nephrology Dialysis Transplantation*, 30(8), 1244-1249, <https://doi.org/10.1093/ndt/gfv071>
- Westfall, P. y Henning, K.S. (2013). *Understanding Advanced Statistical Methods*. Nueva York, Estados Unidos. CRC Press.
- Wickham, H. (2015). *R Packages: Organize, Test, Document, and Share Your Code*. California, Estados Unidos. O'Reilly Media.

- Wickham, H. (2018). scales: Scale Functions for Visualization. Paquete de R versión 1.0.0. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/scales/scales.pdf>
- Wickham, H. y Bryan, J. (2019). readxl: Read Excel Files. Paquete de R versión 1.3.1. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/readxl/readxl.pdf>
- Wickham, H., Chang, W., Henry, L., Lin, T., Takahashi, K., Wilke, C. y Woo, K. (2019). ggplot2: Create Elegant Data Visualizations Using the Grammar of Graphics. . Paquete de R versión 3.1.1. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf>
- Wickham, H. y Henry, L. (2019). tidyr: Easily Tidy Data with “spread()” and “gather()” Functions. Paquete de R version 0.8.3. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/tidyr/tidyr.pdf>
- Willard, W. (2013). HTML: Begginer’s Guide (5ta Ed.). Nueva York, Estado Unidos. McGraw Hill Professional.
- Winkelmann, R. (2013). Econometric Analysis of Count Data (2da Ed.). Berlín, Alemania. Springer Science & Business Media
- Xie, Y., Cheng, J. Tan, X. (2019). DT: A Wrapper of the JavaScript Library “DataTables”. Paquete de R versión 0.6. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/DT/DT.pdf>
- Zeileis, A. y Lumley, T, (2019). sandwich: Robust Covariance Matrix Estimators. Paquete de R version 2.5-1. Recuperado el 11 de Mayo de 2019, desde: <https://cran.r-project.org/web/packages/sandwich/sandwich.pdf>
- Zhao, Z., Sun, Wang, C., Li, P., Wang, W., Ye, J., Gu, X., Wang, X., Shen, S., Zhi, D., Lu, Z., Ye, R., Cheng, R., Xi, L., Zheng, Z., Zhang, M., Luo, F. (2014). Rapidly rising incidence of childhood type 1 diabetes in Chinese population: epidemiology in Shanghai during 1997–2011. *Acta Diabetológica*, 51(6), 947-953.
- Zoni, A. (2018). Epidemiología de las lesiones atendidas en atención primaria en la Comunidad de Madrid (tesis doctoral). Universidad Complutense de Madrid. Madrid, España.
- Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A. y Smith, G.M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. Nueva York, Estados Unidos. Springer.

ANEXOS

ANEXO 1: Sección: Actualizar Datos de la interfaz web.

The screenshot displays a web application interface for 'Intoxicaciones CR'. The interface is divided into a dark sidebar on the left and a main content area on the right. The sidebar contains three menu items: 'Actualizar Datos' (highlighted with a right-pointing arrow), 'Funnel Plot' (with a downward-pointing triangle), and 'Regresión de Poisson' (with a chi-squared symbol). The main content area has a blue header with a hamburger menu icon. Below the header, the text 'Seleccione:' is followed by two file selection sections. The first section is titled 'Datos Intoxicaciones (formato *.xls o *.xlsx):' and contains a 'Browse...' button and a text box with 'No file selected'. The second section is titled 'Datos INEC (formato *.xls o *.xlsx):' and also contains a 'Browse...' button and a text box with 'No file selected'. At the bottom of the main content area, there is a button labeled 'Actualizar datos'.

ANEXO 2: Sección: Funnel Plot de la interfaz web.

Intoxicaciones CR

Actualizar Datos

Funnel Plot

Resultados por # de habitantes:
100.000

Provincia:

Cantón:

Fecha de Consulta:
17/01/01 a 17/12/31

Edad:

Sexo:
 All
 01.Masculino
 02.Femenino
 03.No anotado

Tóxico 1:

CodTox1:

Causa:

Ruta:

Severidad:

911:

Reset Filters

x² Regresión de Poisson

Funnel Plot

Detalles

Show 10 entries Search:

	No_Consulta	Provincia	Canton	Poblacion_Total	No_Intoxicaciones	Fecha_de_Consulta	Edad_anos	Sexo	Toxico_1	CodTox_1
1	8789	Guanacaste	Abangares	18039	21	2017-09-11	55	02.Femenino	FLUMETRINA	08.Plaguici Uso veterir
2	11731	Guanacaste	Abangares	18039	21	2017-12-05	30	01.Masculino	LICOR	13.Abuso
3	5070	Guanacaste	Abangares	18039	21	2017-05-29	54	01.Masculino	DICLORVOS	08.Plaguici Uso veterir
4	1114	Guanacaste	Abangares	18039	21	2017-02-03	16	02.Femenino	ACETAMINOFEN	01.Medicar
5	2474	Guanacaste	Abangares	18039	21	2017-03-16	40	02.Femenino	HIPOCLORITO DE SODIO	04.Hogar y Recreacion
6	5244	Guanacaste	Abangares	18039	21	2017-06-03	23	01.Masculino	IMIDACLOPRID	07.Plaguici Uso agrico
7	9639	Guanacaste	Abangares	18039	21	2017-10-07	23	02.Femenino	HIPOCLORITO DE SODIO	04.Hogar y Recreacion
8	7173	Guanacaste	Abangares	18039	21	2017-07-27	11	01.Masculino	ABEJAS	18.Animale
9	5882	Guanacaste	Abangares	18039	21	2017-06-21	24	01.Masculino	SERPIENTES DESCONOCIDAS	18.Animale
10	11503	Guanacaste	Abangares	18039	21	2017-11-29	99	01.Masculino	VALPROATO	01.Medicar

Showing 1 to 10 of 9,635 entries Previous 1 2 3 4 5 ... 964 Next

Detalles secundarios

ANEXO 3: Sección: Regresión de Poisson, sub sección: Configuración de la interfaz web.

Intoxicaciones CR

Actualizar Datos
Funnel Plot
Regresión de Poisson

Filtros:

Variable 1:
Variable 2:
Variable 3:
Reset Filters

Poisson Configuración Resultados

1. Seleccione la o las variables predictoras:

Variable 1: Canton
Variable 2: -
Variable 3: -

2. Defina el No. de intoxicaciones para agrupar en la categoría 'Otros'.

Categoría 1: 0

3. Seleccione el nivel de referencia:

Valor de referencia 1:
Valor de referencia 2: -
Valor de referencia 3: -

Permitir Interacciones: No

ANEXO 4: Sección: Regresión de Poisson, sub sección: Resultados de la interfaz web.

Intoxicaciones CR
☰

Poisson Configuración Resultados

163.785

AIC

0.946

R² (McFadden)

Posibles interpretaciones:

- El **AIC** es una medida relativa para comparar modelos. Se considera un mejor modelo si presenta un menor AIC con respecto a otro modelo.
- Un valor de **R² (McFadden)** cercano a 1 indica que el modelo explica cerca del 100% de la variabilidad de los datos con respecto al modelo nulo (sin predictores).

Prueba de Sobredispersión

Show 10 entries Search:

	Dispersión	z	Valor p
Func. Lineal	-0.174	-0.591	0.723
Func. Cuadrática	-0.001	-0.984	0.837

Showing 1 to 2 of 2 entries Previous 1 Next

Posibles interpretaciones:

- Dispersión > 0 indica sobredispersión (varianza > media) en los datos si $P < 0.05$.
- Se asume sobredispersión con un valor de dispersión > 0 en la función lineal o cuadrática y un $P < 0.05$.

Factor de Inflación de la Varianza (FIV)

Show 10 entries Search:

	GFIV	GL	GFIV^(1/2GL)	GFIV^(1/GL)
var1	1	3	1	1
var2	1	4	1	1

Showing 1 to 2 of 2 entries Previous 1 Next

Posibles interpretaciones:

- $GFIV^{(1/GL)} = FIV$
- El FIV solo puede ser calculado si se han escogido 2 o más variables en el paso 1.
- Indica si un estimador (variable seleccionada en el paso 1) no es confiable.
- Valores de FIV > 10 sugieren que la variable debe de ser removida.

Tasa de Incidencia

Show 20 entries Search:

	IRR	Errores Robustos	LI	LS	Valor p	Significativo
(Intercept)	783.799	15.003	754.937	813.763	0	Si
var1Alajuela	0.549	0.021	0.509	0.592	0	Si
var1Cartago	0.283	0.015	0.256	0.313	0	Si
var1Heredia	0.278	0.008	0.263	0.294	0	Si
var2b. 4-6	0.279	0.011	0.259	0.301	0	Si
var2c. 7-12	0.252	0.005	0.242	0.263	0	Si
var2d. 13-15	0.148	0.011	0.129	0.171	0	Si
var2e. 16-19	0.265	0.015	0.236	0.297	0	Si

Showing 1 to 8 of 8 entries Previous 1 Next

Posibles interpretaciones:

- Si el valor del **IRR** es **menor a 1** se considera que hay una menor cantidad de intoxicaciones con respecto al nivel de referencia (elegido en la sección anterior).
 - El porcentaje de disminución de intoxicaciones es el resultado de $(1-IRR)*100$.
- Si el valor **IRR** es **mayor a 1** se considera que hay una mayor cantidad de intoxicaciones con respecto al nivel de referencia (elegido en la sección anterior).
 - El porcentaje de aumento de intoxicaciones es el resultado de $(IRR-1)*100$.
 - Si el valor es de 2 o mayor, se puede leer como 2, 3, 4, ..., n veces más intoxicaciones con respecto al nivel de referencia.

Poisson - resultados crudos +

Actualizar Datos
Funnel Plot
x² Regresión de Poisson

Filtrros:

Variable 1:
Alajuela Cartago Heredia
San Jose
Variable 2:
a. 0-3 b. 4-6 c. 7-12
d. 13-15 e. 16-19
Variable 3:

Reset Filters

ANEXO 5: Código para la aplicación creada para el análisis de intoxicaciones en Costa Rica.

```
#install.packages("shiny")
#install.packages("shinydashboard")
#install.packages("tidyr")
#install.packages("readxl")
#install.packages('ggplot2')
#install.packages("scales")
#install.packages("plotly")
#install.packages("DT")
#install.packages("sandwich")
#install.packages("modEvA")
#install.packages("car")
#install.packages("AER")

#-----
#Required packages
#-----
library(shiny)
library(shinydashboard)
library(plotly)
library(readxl)
library(ggplot2)
library(tidyr)
library(DT)
library(scales)
library(sandwich)
library(modEvA)
library(car)
library(AER)

options(scipen=999, shiny.maxRequestSize=100*1024^2) # Para evitar notacion cientifica y problemas por exceso de datos

#-----
#ui object
#-----
ui <- dashboardPage(
  #DASHBOARD HEADER
  dashboardHeader(title = "Intoxicaciones CR"),
  #DASHBOARD SIDEBAR
  dashboardSidebar(
    #Sidebar content
    sidebarMenu(id = "sidebarmenu",
      menuItem("Actualizar Datos", tabName = "datos", icon = icon("sign-in")),
      menuItem("Funnel Plot", tabName = "funnelinfo", icon = icon("filter")),
      conditionalPanel("input.sidebarmenu == 'funnelinfo'",
        #actionButton(inputId = 'reset', label = 'Reset map marking'),
        selectInput('habitantes', label = 'Resultados por # de habitantes:', choices = c('1.000.000' =
1000000, '100.000'=100000, '10.000'=10000, '1.000'=1000), selected = 100000),
        selectInput('provincia', label = 'Provincia:', choices = character(0), multiple = T),
        selectInput('canton', label = 'Cantón:', choices = character(0), multiple = T),
        dateRangeInput('fechaConsulta', label = 'Fecha de Consulta:',
          start = "2012-01-01",
          end = "2012-12-31",
          min = "2012-01-01",
          max = "2012-12-31",
          format = "yy/mm/dd",
          language = "es",
          separator = " a "),
        sliderInput('edad', label = 'Edad:',
          min = -1,
          max = 100,
          value = c(-1,100)),
        radioButtons('sexo', label = 'Sexo:', choices = list( a = 1, b = 2, "All"), selected = "All"),
        selectInput('tox1', label = 'Tóxico 1:', choices = character(0), multiple = T),
        selectInput('codtox1', label = 'CodTox1:', choices = character(0), multiple = T),
        selectInput('causa', label = 'Causa:', choices = character(0), multiple = T),
        selectInput('ruta', label = 'Ruta:', choices = character(0), multiple = T),
        selectInput('sev', label = 'Severidad:', choices = character(0), multiple = T),
        selectInput('nuevell', label = '911:', choices = character(0), multiple = T),
        actionButton('reset', label = 'Reset Filters')
      ),
    menuItem("Regresión de Poisson", tabName = "regresion", icon = icon("superscript")),
    conditionalPanel("input.sidebarmenu == 'regresion'",
      h5("Filtros:"),
      selectInput('varFilter1', label = 'Variable 1:', choices = character(0), multiple = T),
      selectInput('varFilter2', label = 'Variable 2:', choices = character(0), multiple = T),
      selectInput('varFilter3', label = 'Variable 3:', choices = character(0), multiple = T),
      actionButton('resetPoi', label = 'Reset Filters')
    )
  )
),
  dashboardBody(
```

```

tabItems(
  #Tab para actualizar datos
  tabItem(tabName = "datos",
    tags$h5(tags$strong("Seleccione:")),
    fileInput(inputId = 'centro', label = 'Datos Intoxicaciones (formato *.xls o *.xlsx):',
      accept = c(".xlsx", ".xls")),
    fileInput(inputId = 'inec', label = 'Datos INEC (formato *.xls o *.xlsx):',
      accept = c(".xlsx", ".xls")),
    actionButton(inputId = 'actualizar', label = 'Actualizar datos')
  ),
  #Tab para analizar el funnel plot y detalles
  tabItem(tabName = "funnelinfo",
    #Funnel
    fluidRow(
      box(title = "Funnel Plot", width = 12, plotlyOutput("funnel")
    )
  ),
  #Tabla de detalles
  fluidRow(
    box(
      title = "Detalles", width = 12, status = "primary",
      collapsible = TRUE, dataTableOutput("detallesfun")
    )
  ),
  #Tabla de detalles 2
  fluidRow(
    box(
      title = "Detalles secundarios", width = 12, status = "primary",
      collapsible = TRUE, collapsed = TRUE, dataTableOutput("detallesfun2")
    )
  )
),
#Tab para analizar Poisson
tabItem(tabName = "regresion",
  navbarPage(title = "Poisson",
    tabPanel("Configuracion",
      fluidRow(
        column(width = 4,
          tags$h4(tags$strong("1. Seleccione la o las variables predictoras:")),
          br(),
          selectInput("var1", "Variable 1:", choices = sort(c("Canton", "Provincia",
            "RangoEtario", "Sexo", "Toxicol", "Codtox1", "Causa", "Ruta", "Severidad", "NueveOnce", "Entidades"))),
          selectInput("var2", "Variable 2:", choices = c("-", sort(c("Canton", "Provincia",
            "RangoEtario", "Sexo", "Toxicol", "Codtox1", "Causa", "Ruta", "Severidad", "NueveOnce", "Entidades")))),
          selectInput("var3", "Variable 3:", choices = c("-", sort(c("Canton", "Provincia",
            "RangoEtario", "Sexo", "Toxicol", "Codtox1", "Causa", "Ruta", "Severidad", "NueveOnce", "Entidades"))))
        ),
        column(width = 4,
          tags$h4(tags$strong("2. Defina el No. de intoxicaciones para agrupar en la categoria
            'Otros'.")),
          conditionalPanel(
            condition = "input.var1 == 'Canton' | input.var1 == 'Toxicol' | input.var1 ==
            'Codtox1' | input.var1 == 'Causa' | input.var1 == 'Ruta' | input.var1 == 'Entidades'",
            numericInput("num1", "Categoria 1:", value = 0, min = 0, max = 1000, step = 1)
          ),
          conditionalPanel(condition = "input.var1 == 'Provincia' | input.var1 == 'Sexo' |
            input.var1 == 'RangoEtario' | input.var1 == 'Severidad' | input.var1 == 'NueveOnce'",
            br(),
            br(),
            br(),
            br()
          ),
          conditionalPanel(
            condition = "input.var2 == 'Canton' | input.var2 == 'Toxicol' | input.var2 ==
            'Codtox1' | input.var2 == 'Causa' | input.var2 == 'Ruta' | input.var2 == 'Entidades'",
            numericInput("num2", "Categoria 2:", value = 0, min = 0, max = 1000, step = 1)
          ),
          conditionalPanel(condition = "input.var1 == 'Provincia' | input.var1 == 'Sexo' |
            input.var1 == 'RangoEtario' | input.var1 == 'Severidad' | input.var1 == 'NueveOnce'",
            br(),
            br()
          ),
          conditionalPanel(
            condition = "input.var3 == 'Canton' | input.var3 == 'Toxicol' | input.var3 ==
            'Codtox1' | input.var3 == 'Causa' | input.var3 == 'Ruta' | input.var3 == 'Entidades'",
            numericInput("num3", "Categoria 3:", value = 0, min = 0, max = 1000, step = 1)
          )
        ),
        column(width = 4,
          tags$h4(tags$strong("3. Seleccione el nivel de referencia:")),
          br(),
          selectInput("ref1", "Valor de referencia 1:", choices = character(0)),
          selectInput("ref2", "Valor de referencia 2:", choices = character(0)),
          selectInput("ref3", "Valor de referencia 3:", choices = character(0))
        )
      )
    )
  )
),
br()

```



```

#-----#
#server object
#-----#
server <- function(input, output, session) {
  #-----#
  #Reactive Functions
  #-----#

  #----#
  #Cargar datos
  #----#
  paths <- eventReactive(input$actualizar,{
    #Salva el path de los datos del Centro y de censo poblacion (inec) en las variables dircentro y dirpoblacion.
    ###centro <- read.csv(dircentro)
    dircentro <- input$centro

    dirpoblacion <- input$inec

    updateTabItems(session = session, inputId = "sidebarmenu", 'funnelinfo')

    list(dircentro = dircentro, dirpoblacion = dirpoblacion)
  })

  #----#
  #Creacion de UNA tabla con datos del centro y del censo (INEC) = centroCenso
  #----#
  centroCensoTableBase <- reactive({
    dircentro <- paths()$dircentro$datapath
    dirpoblacion <- paths()$dirpoblacion$datapath

    ###Cargar datos del Centro

    #Salva los datos del Centro en la variable centro.
    centro <- as.data.frame(read_xlsx(dircentro, sheet = 1), stringsAsFactors = FALSE)

    #Se remueven los datos de animales
    centro <- centro[!grepl("ANI", centro$TipoIntox, ignore.case = TRUE),]

    #Se remueven los datos de asesoria
    centro <- centro[!grepl("No", centro$Intoxicacion, ignore.case = TRUE),]

    #Se renombra la columna de Codtox en caso de que no tenga el formato de 2017 y 2018
    if("CodTox1" %in% colnames(centro)){
      colnames(centro)[colnames(centro)=="CodTox1"] <- "Codtox1"
      colnames(centro)[colnames(centro)=="CodTox2"] <- "Codtox2"
      colnames(centro)[colnames(centro)=="CodTox3"] <- "Codtox3"
      colnames(centro)[colnames(centro)=="CodTox4"] <- "Codtox4"
      colnames(centro)[colnames(centro)=="CodTox5"] <- "Codtox5"
      colnames(centro)[colnames(centro)=="CodTox6"] <- "Codtox6"
    }

    #Se limpian los datos de provincia y canton separandolos del ID
    centro <- separate(data=centro,
                      col=Canton,
                      into=c("id_Canton", "Canton"),
                      extra="merge")

    centro <- separate(data=centro,
                      col=Provincia,
                      into=c("id_Provincia", "Provincia"),
                      extra="merge")

    centro$Fecha <- as.Date(centro$Fecha)

    centro$NueveOnce[is.na(centro$NueveOnce) | centro$NueveOnce == 0 | centro$NueveOnce == "02.No"] <- "No"
    centro$NueveOnce[centro$NueveOnce == 1 | centro$NueveOnce == "01.Si"] <- "Si"

    ###Cargar datos del INEC

    #Salva los datos del INEC en la variable censo
    censo <- as.data.frame(read_excel(dirpoblacion,
                                     sheet=3), stringsAsFactors = FALSE)

    #Join de la tabla del centro con la tabla de población cantonal
    centroCenso <- merge(x = centro,
                        y = censo,
                        by.x = c("Canton", "Provincia"),
                        by.y = c("Canton", "Provincia"),
                        all.x = TRUE
    )

    list(censo = censo, centroCenso = centroCenso, centro = centro)
  })

  #-----#

```

```

#FUNNEL PLOT
#-----#

#----#
#Actualizar inputs de Canton y Fecha de Consulta para Funnel
#----#

observe({
  centroCenso2 <- centroCensoTableBase()$centroCenso

  #UpdateInput: Canton
  updateSelectInput(session, "canton",
                    #label = "Cantón:",
                    choices = sort(unique(na.omit(centroCenso2$Canton))),
                    selected = character(0))

  #UpdateInput: Fecha
  updateDateRangeInput(session, 'fechaConsulta',
                       label = 'Fecha de Consulta:',
                       start = min(centroCenso2$Fecha, na.rm = TRUE),
                       end = max(centroCenso2$Fecha, na.rm = TRUE),
                       min = min(centroCenso2$Fecha, na.rm = TRUE),
                       max = max(centroCenso2$Fecha, na.rm = TRUE))
})

#----#
#Filtro tablas funnel
#----#
centroCensoTable <- reactive ({

  centroCenso2 <- centroCensoTableBase()$centroCenso
  censo2 <- centroCensoTableBase()$censo

  #Evaluar filtro de: provincia
  if(is.null(input$provincia)){
    provres <- unique(centroCenso2$Provincia)
    centroCenso <- subset(centroCenso2, Provincia %in% provres)
    censo <- subset(censo2, Provincia %in% provres)
    updateSelectInput(session, "provincia",
                      #label = "Provincia:",
                      choices = sort(unique(censo2$Provincia)),
                      selected = character(0))
  } else {
    provres <- input$provincia
    centroCenso <- subset(centroCenso2, Provincia %in% provres)
    censo <- subset(censo2, Provincia %in% provres)
  }

  #NO HAY FILTRO DE CANTON (Esto porque quedan pocos datos para que el funnel tenga sentido)

  #Evaluar filtro de: fechaConsulta
  minfechares <- input$fechaConsulta[1]
  maxfechares <- input$fechaConsulta[2]
  centroCenso <- subset(centroCenso, Fecha >= minfechares & Fecha <= maxfechares)

  #Evaluar filtro de: edad
  if(input$edad[1] == min(as.numeric(na.omit(centroCenso$Edadanos))) & input$edad[2] ==
max(as.numeric(na.omit(centroCenso$Edadanos)))) {
    minedadres <- min(as.numeric(na.omit(centroCenso$Edadanos)))
    maxedadres <- max(as.numeric(na.omit(centroCenso$Edadanos)))
    centroCenso <- subset(centroCenso, Edadanos >= minedadres & Edadanos <= maxedadres)
    updateSliderInput(session, 'edad',
                      label = 'Edad:',
                      value =
c(min(as.numeric(na.omit(centroCenso2$Edadanos))),max(as.numeric(na.omit(centroCenso2$Edadanos))))),
                      min = minedadres,
                      max = maxedadres)
  } else {
    minedadres <- input$edad[1]
    maxedadres <- input$edad[2]
    centroCenso <- subset(centroCenso, Edadanos >= minedadres & Edadanos <= maxedadres)
    updateSliderInput(session, 'edad',
                      label = 'Edad:',
                      value = c(minedadres,maxedadres),
                      min = min(as.numeric(na.omit(centroCenso2$Edadanos))),
                      max = max(as.numeric(na.omit(centroCenso2$Edadanos))))
  }

  #Evaluar filtro de: sexo
  if(input$sexo == "All"){
    sexres <- as.character(unique(centroCenso$Sexo))
    centroCenso <- subset(centroCenso, Sexo %in% sexres)
    updateRadioButtons(session, "sexo",
                       #label = "Sexo:",

```

```

                choices = c("All", sort(unique(centroCenso$Sexo))),
                selected = "All")
} else {
  sexres <- input$ssexo
  centroCenso <- subset(centroCenso, Sexo %in% sexres)
}

#Evaluar filtro de: tox1
if(is.null(input$tox1)){
  tox1res <- as.character(unique(centroCenso$Toxicol))
  centroCenso <- subset(centroCenso, Toxicol %in% tox1res)
  updateSelectInput(session, "tox1",
                    #label = "Toxicol:",
                    choices = sort(unique(na.omit(centroCenso$Toxicol))),
                    selected = character(0))
}
} else {
  tox1res <- input$tox1
  centroCenso <- subset(centroCenso, Toxicol %in% tox1res)
}

#Evaluar filtro de: codtox1
if(is.null(input$codtox1)){
  codtox1res <- as.character(unique(centroCenso$Codtox1))
  centroCenso <- subset(centroCenso, centroCenso$Codtox1 %in% codtox1res)
  updateSelectInput(session, "codtox1",
                    #label = "Codtox1:",
                    choices = sort(unique(na.omit(centroCenso$Codtox1))),
                    selected = character(0))
}
} else {
  codtox1res <- input$codtox1
  centroCenso <- subset(centroCenso, centroCenso$Codtox1 %in% codtox1res)
}

#Evaluar filtro de: causa
if(is.null(input$causa)){
  caures <- as.character(unique(centroCenso$Causa))
  centroCenso <- subset(centroCenso, centroCenso$Causa %in% caures)
  updateSelectInput(session, "causa",
                    #label = "Causa:",
                    choices = sort(unique(na.omit(centroCenso$Causa))),
                    selected = character(0))
}
} else {
  caures <- input$causa
  centroCenso <- subset(centroCenso, centroCenso$Causa %in% caures)
}

#Evaluar filtro de: ruta
if(is.null(input$ruta)){
  rutres <- as.character(unique(centroCenso$Ruta))
  centroCenso <- subset(centroCenso, centroCenso$Ruta %in% rutres)
  updateSelectInput(session, "ruta",
                    #label = "Ruta:",
                    choices = sort(unique(na.omit(centroCenso$Ruta))),
                    selected = character(0))
}
} else {
  rutres <- input$ruta
  centroCenso <- subset(centroCenso, centroCenso$Ruta %in% rutres)
}

#Evaluar filtro de: severidad
if(is.null(input$sev)){
  sevres <- as.character(unique(centroCenso$Severidad))
  centroCenso <- subset(centroCenso, centroCenso$Severidad %in% sevres)
  updateSelectInput(session, "sev",
                    #label = "Severidad:",
                    choices = sort(unique(na.omit(centroCenso$Severidad))),
                    selected = character(0))
}
} else {
  sevres <- input$sev
  centroCenso <- subset(centroCenso, centroCenso$Severidad %in% sevres)
}

#Evaluar filtro de: nuevell
if(is.null(input$nuevell)){
  nueres <- as.character(unique(centroCenso$NueveOnce))
  centroCenso <- subset(centroCenso, centroCenso$NueveOnce %in% nueres)
  updateSelectInput(session, "nuevell",
                    #label = "NueveOnce:",
                    choices = sort(unique(na.omit(centroCenso$NueveOnce))),
                    selected = character(0))
}
}

```

```

    )
  } else {
    nueres <- input$nueve11
    centroCenso <- subset(centroCenso, centroCenso$NueveOnce %in% nueres)
  }

  list(censo = censo, centroCenso = centroCenso)
})

#----#
#Reseteo de filtros y marking del funnel (Luego del doble click)
#----#
observeEvent(input$reset,{
  observe({
    centroCenso <- centroCensoTableBase()$centroCenso
    censo <- centroCensoTableBase()$censo

    #Reseteo filtro: Provincia
    provres <- unique(centroCenso$Provincia)
    centroCenso <- subset(centroCenso, Provincia %in% provres)
    censo <- subset(censo, Provincia %in% provres)
    updateSelectInput(session, "provincia",
                      #label = "Provincia:",
                      choices = sort(unique(censo$Provincia)),
                      selected = character(0))

    #Reseteo filtro: canton
    updateSelectInput(session, "canton",
                      #label = "Cantón:",
                      choices = sort(unique(na.omit(centroCenso$Canton))),
                      selected = character(0))

    #Reseteo filtro: fechaConsulta
    updateDateRangeInput(session, 'fechaConsulta',
                          label = 'Fecha de Consulta:',
                          start = min(centroCenso$Fecha),
                          end = max(centroCenso$Fecha),
                          min = min(centroCenso$Fecha),
                          max = max(centroCenso$Fecha))

    #Reseteo filtro: edad
    minedadres <- min(as.numeric(na.omit(centroCenso$Edadanos)))
    maxedadres <- max(as.numeric(na.omit(centroCenso$Edadanos)))
    centroCenso <- subset(centroCenso, centroCenso$Edadanos >= minedadres & centroCenso$Edadanos <= maxedadres)
    updateSliderInput(session, 'edad',
                      label = 'Edad:',
                      value =
c(min(as.numeric(na.omit(centroCenso$Edadanos))),max(as.numeric(na.omit(centroCenso$Edadanos))))),
                      min = minedadres,
                      max = maxedadres)

    #Reseteo filtro: sexo
    sexres <- as.character(unique(centroCenso$Sexo))
    centroCenso <- subset(centroCenso, Sexo %in% sexres)
    updateRadioButtons(session, "sexo",
                       #label = "Sexo:",
                       choices = c("All", sort(unique(centroCenso$Sexo))),
                       selected = "All")

    #Reseteo filtro: tox1
    tox1res <- as.character(unique(centroCenso$Toxicol))
    centroCenso <- subset(centroCenso, centroCenso$Toxicol %in% tox1res)
    updateSelectInput(session, "tox1",
                      #label = "Toxicol:",
                      choices = sort(unique(na.omit(centroCenso$Toxicol))),
                      selected = character(0))

    #Reseteo filtro: codtox1
    codtox1res <- as.character(unique(centroCenso$Codtox1))
    centroCenso <- subset(centroCenso, centroCenso$Codtox1 %in% codtox1res)
    updateSelectInput(session, "codtox1",
                      #label = "Codtox1:",
                      choices = sort(unique(na.omit(centroCenso$Codtox1))),
                      selected = character(0))

    #Reseteo filtro: causa
    caures <- as.character(unique(centroCenso$Causa))
    centroCenso <- subset(centroCenso, centroCenso$Causa %in% caures)
    updateSelectInput(session, "causa",
                      #label = "Causa:",
                      choices = sort(unique(na.omit(centroCenso$Causa))),
                      selected = character(0))

    #Reseteo filtro: ruta
    rutres <- as.character(unique(centroCenso$Ruta))

```

```

centroCenso <- subset(centroCenso, centroCenso$Ruta %in% rutres)
updateSelectInput(session, "ruta",
  #Label = "Ruta:",
  choices = sort(unique(na.omit(centroCenso$Ruta))),
  selected = character(0))

#Resetear filtro: sev
sevres <- as.character(unique(centroCenso$Severidad))
centroCenso <- subset(centroCenso, centroCenso$Severidad %in% sevres)
updateSelectInput(session, "sev",
  #Label = "Severidad:",
  choices = sort(unique(na.omit(centroCenso$Severidad))),
  selected = character(0))

#Resetear filtro: nueve11
nueres <- as.character(unique(centroCenso$NueveOnce))
centroCenso <- subset(centroCenso, centroCenso$NueveOnce %in% nueres)
updateSelectInput(session, "nueve11",
  #Label = "NueveOnce:",
  choices = unique(na.omit(centroCenso$NueveOnce)),
  selected = character(0))
})
})

#----#
#Creacion de la tabla requerida para el funnel y tabla de detalles = funnelTable
#----#
funnelTableTable <- reactive({
  centroCenso <- centroCensoTable()$centroCenso
  censo <- centroCensoTable()$censo

  #####Funnel plot o grafico de embudo

  #Preparación de las variables
  #Eje y: total de intoxicaciones por cantón
  #NoConsulta (fuente: Centro)
  #Cantón (fuente: Centro)
  #Tóxiol (fuente: Centro)
  #DIVIDIDO por la población total por cantón (fuente: censo)
  #Eje x: población total por cantón (fuente: censo)
  #Limites de control inferior y superior
  #Promedio de la tasa de intoxicaciones

  #total de intoxicaciones por cantón
  intoxicacionCanton <- aggregate(
    centroCenso$NoConsulta,
    by=list(Canton=centroCenso$Canton),
    FUN=length)

  #Renombramos la columna X
  colnames(intoxicacionCanton)[colnames(intoxicacionCanton)=="x"] <- "NumIntoxicacionesCanton"

  #Join de la tabla con intoxicacionCanton y censo
  #total por cantón
  funnelTable <- merge(x = censo,
    y = intoxicacionCanton,
    by = "Canton",
    all.x = TRUE
  )

  #Reemplazo todos los NA en la columna de Numero de Intoxicaciones con un 0 para que el punto de todos los cantones aparezca
  aunque no tengan intoxicaciones.
  funnelTable[is.na(funnelTable$NumIntoxicacionesCanton), 'NumIntoxicacionesCanton'] <- 0

  #División del total de intoxicaciones por cantón
  #(intoxicacionCanton=X=NumIntoxicacionesCanton dentro del funnelTable)
  #entre el total de la población en el cantón (eje yfunnel = tasa)

  if(toupper(substr(input$sexo, nchar(input$sexo)-3, nchar(input$sexo))) == "LINO"){
    funnelTable <- funnelTable %>%
      mutate(Tasa = NumIntoxicacionesCanton / Hombres)
  } else if (toupper(substr(input$sexo, nchar(input$sexo)-3, nchar(input$sexo))) == "NINO"){
    funnelTable <- funnelTable %>%
      mutate(Tasa = NumIntoxicacionesCanton / Mujeres)
  } else {
    funnelTable <- funnelTable %>%
      mutate(Tasa = NumIntoxicacionesCanton / Total)
  }

  #Renombramos las columna Total con ejex
  if(toupper(substr(input$sexo, nchar(input$sexo)-3, nchar(input$sexo))) == "LINO"){
    colnames(funnelTable)[colnames(funnelTable)=="Hombres"] <- "Poblacion_por_Canton"
  } else if (toupper(substr(input$sexo, nchar(input$sexo)-3, nchar(input$sexo))) == "NINO"){
    colnames(funnelTable)[colnames(funnelTable)=="Mujeres"] <- "Poblacion_por_Canton"
  } else {

```

```

    colnames(funnelTable)[colnames(funnelTable)=="Total"] <- "Poblacion_por_Canton"
  }
}

funnelTable
})

#----#
#Calculos requeridos para el funnel = m, nseq, limInf95, limInf99, limSup95, limSup99, col
#----#
funnelReq <- reactive({
  funnelTable <- funnelTableTable()

  #Columnas requeridas
  #Promedio de la tasa eliminando los nulos
  m <- mean(funnelTable$Tasa, na.rm = TRUE)
  #Se generan 1000 puntos entre el valor mimino y maximo del total
  #poblacional por canton para los limites de control
  nseq <- seq(min(na.omit(funnelTable$Poblacion_por_Canton)),max(na.omit(funnelTable$Poblacion_por_Canton)),length.out =
1000)
  #Cada punto se multiplica por el promedio de la tasa de intoxicaciones
  y <- m*nseq #total de intoxicaciones tomando en cuenta la tasa

  #Se calculan lo limites sup e inf (95 y 99% de confianza) considerando los valores de z
  #y = numero de intoxicaciones por canton
  #z = se debe de calcular como:
  #0.05 para un 95% de confianza, 0.01 para un 99% de confianza
  #100*(1-(0.05/2))
  #qnorm(0.975)
  #qnorm(0.025)

  #100*(1-(0.01/2))
  #qnorm(0.995)
  #qnorm(0.005)
  ymin95 = y * (1 - (1/(9*y)) - (qnorm(0.975)/(3*sqrt(y))))^3
  ymin99 = y * (1 - (1/(9*y)) - (qnorm(0.995)/(3*sqrt(y))))^3
  ymax95 = (y+1) * (1-(1/(9*(y+1)))+(qnorm(0.975)/(3*sqrt(y+1))))^3
  ymax99 = (y+1) * (1-(1/(9*(y+1)))+(qnorm(0.995)/(3*sqrt(y+1))))^3

  limInf95 = ymin95/nseq
  limInf99 = ymin99/nseq
  limSup95 = ymax95/nseq
  limSup99 = ymax99/nseq

  dfCI <- data.frame(nseq,limInf95, limInf99, limSup95, limSup99)

  #Se realizan las interpolaciones de los valores reales con los limites de control para det rojo/amarillo/verde
  interInf95 <- approx(x=dfCI$nseq,
                      y=dfCI$limInf95,
                      xout = na.omit(funnelTable$Poblacion_por_Canton))

  interInf99 <- approx(x=dfCI$nseq,
                      y=dfCI$limInf99,
                      xout = na.omit(funnelTable$Poblacion_por_Canton))

  interSup95 <- approx(x=dfCI$nseq,
                      y=dfCI$limSup95,
                      xout = na.omit(funnelTable$Poblacion_por_Canton))

  interSup99 <- approx(x=dfCI$nseq,
                      y=dfCI$limSup99,
                      xout = na.omit(funnelTable$Poblacion_por_Canton))

  #Agregar una columna con la indicacion de color por Canton
  funnelTable['col'] <- ""

  #Se crea un vector con la indicacion de color para cada punto del funnel.
  for (i in 1:length(na.omit(funnelTable$Tasa))){
    if (#(na.omit(funnelTable$Tasa)[i] > interInf99$y[i] & na.omit(funnelTable$Tasa)[i] < interInf95$y[i]) |
        (na.omit(funnelTable$Tasa)[i] > interSup95$y[i] & na.omit(funnelTable$Tasa)[i] < interSup99$y[i])){
      funnelTable$col[i] <- 'darkgoldenrod1'
    } else if (#na.omit(funnelTable$Tasa)[i] < interInf99$y[i] |
              na.omit(funnelTable$Tasa)[i] > interSup99$y[i]) {
      funnelTable$col[i] <- 'firebrick'
    } else {
      funnelTable$col[i] <- 'chartreuse4'
    }
  }
}

#Cambio del color del punto de acuerdo con el canton
for(i in 1:length(funnelTable$Canton)){
  if(funnelTable$Canton[i] %in% input$canton){
    funnelTable$col[i] <- 'black'
  }
}
}

```

```

#Variable que indica por cada cuátos habitantes se están dando los resultados
hab <- as.numeric(gsub(',', '', format(input$habitantes, scientific = FALSE, big.mark = ",")))

key = unique(funnelTable$Canton)

funnelPlot <-
ggplot(na.omit(funnelTable), aes(x=Poblacion_por_Canton, y=Tasa*hab)) +
geom_point(aes(Provincia = Provincia, Canton = Canton, key = key), color = funnelTable$col) +
xlab('Población por Cantón') +
ylab(paste0('Intox por cada ', format(as.numeric(input$habitantes), scientific = FALSE, big.mark = ","), ' hab')) +
ylim(min(na.omit(funnelTable$Tasa))*hab-1, max(na.omit(funnelTable$Tasa))*hab+1) +
geom_hline(yintercept = m*hab) +
#geom_line(aes(x = nseq, y = limInf95*hab), color = 'darkgoldenrod1', data = dfCI) +
#geom_line(aes(x = nseq, y = limInf99*hab), color = 'firebrick', data = dfCI) +
geom_line(aes(x = nseq, y = limSup95*hab), color = 'darkgoldenrod1', data = dfCI) +
geom_line(aes(x = nseq, y = limSup99*hab), color = 'firebrick', data = dfCI) +
scale_x_continuous(labels = comma)

list(m = m, nseq = nseq, limInf95 = limInf95, limInf99 = limInf99, limSup95 = limSup95, limSup99 = limSup99, dfCI = dfCI,
funnelPlot = funnelPlot)
})

#----#
#Calculos requeridos para las tablas de detalles del funnel
#----#
funDetTablePrep <- reactive ({
centroCenso <- centroCensoTable()$centroCenso

#total de intoxicaciones por cantón
intoxicacionCanton <- aggregate(
centroCenso$NoConsulta,
by=list(Canton=centroCenso$Canton),
FUN=length)

#Join de la tabla intoxicacionCanton y centroCenso
#total por cantón
funDetTablePrep <- merge(x = centroCenso,
y = intoxicacionCanton,
by = "Canton",
all.x = TRUE
)

funDetTablePrep
})

#----#
#Tablas de detalles del funnel plot
#----#
funDetTables <- reactive ({
funDetTablePrep <- funDetTablePrep()

keyres <- event_data("plotly_selected")$key

#Evaluar marking del funnelplot
if(is.null(keyres)){
cantres <- as.character(unique(funDetTablePrep$Canton))
funDetTablePrep <- subset(funDetTablePrep, Canton %in% cantres)
} else {
cantres <- keyres
funDetTablePrep <- subset(funDetTablePrep, Canton %in% cantres)
}

funDetails <- data.frame(No_Consulta = funDetTablePrep$NoConsulta,
Provincia = funDetTablePrep$Provincia,
Canton = funDetTablePrep$Canton,
Poblacion_Total = funDetTablePrep$Total,
No_Intoxicaciones = funDetTablePrep$x, #viene del centroCenso con calculo y join,
Fecha_de_Consulta = funDetTablePrep$Fecha,
Edad_anos = funDetTablePrep$Edadanos,
Sexo = funDetTablePrep$Sexo,
Toxico_1 = funDetTablePrep$Toxico1,
CodTox_1 = funDetTablePrep$Codtox1,
Entidades = funDetTablePrep$Entidades,
Causa = funDetTablePrep$Causa,
Ruta = funDetTablePrep$Ruta,
Severidad = funDetTablePrep$Severidad,
NueveOnce = funDetTablePrep$NueveOnce)

funDetails2 <- data.frame(No_Consulta = funDetTablePrep$NoConsulta,
Toxico_1 = funDetTablePrep$Toxico1,
CodTox_1 = funDetTablePrep$Codtox1,
Toxico_2 = funDetTablePrep$Toxico2,
CodTox_2 = funDetTablePrep$Codtox2,
Toxico_3 = funDetTablePrep$Toxico3,

```

```

        CodTox_3 = funDetTablePrep$Codtox3,
        Toxico_4 = funDetTablePrep$Toxico4,
        CodTox_4 = funDetTablePrep$Codtox4,
        Toxico_5 = funDetTablePrep$Toxico5,
        CodTox_5 = funDetTablePrep$Codtox5,
        Toxico_6 = funDetTablePrep$Toxico6,
        CodTox_6 = funDetTablePrep$Codtox6)

    list(funDetails = funDetails, funDetails2 = funDetails2)
  })

#-----
#Render Functions
#-----

#Render Funnel
output$funnel <- renderPlotly({
  funnelPlot <- funnelReq()$funnelPlot
  funnelPlot <- ggplotly(funnelPlot) %>%
    layout(dragmode = "select")
  funnelPlot
})

#Render detalles del Funnel
output$detallesfun <- renderDataTable({
  funDetails <- funDetTables()$funDetails
  datatable(funDetails, options = list(scrollX = TRUE))
})

#Render detalles secundarios del Funnel
output$detallesfun2 <- renderDataTable({
  funDetails2 <- funDetTables()$funDetails2
  datatable(funDetails2, options = list(scrollX = TRUE))
})

#-----#
#POISSON
#-----#

#----#
#Calculo de columna de RangoEtario
#----#
coledad <- reactive({
  centro <- centroCensoTableBase()$centro

  centro$RangoEtario[centro$Edadanos >= 0 & centro$Edadanos <= 3] <- "a. 0-3"
  centro$RangoEtario[centro$Edadanos >= 4 & centro$Edadanos <= 6] <- "b. 4-6"
  centro$RangoEtario[centro$Edadanos >= 7 & centro$Edadanos <= 12] <- "c. 7-12"
  centro$RangoEtario[centro$Edadanos >= 13 & centro$Edadanos <= 15] <- "d. 13-15"
  centro$RangoEtario[centro$Edadanos >= 16 & centro$Edadanos <= 19] <- "e. 16-19"
  centro$RangoEtario[centro$Edadanos >= 20 & centro$Edadanos <= 29] <- "f. 20-29"
  centro$RangoEtario[centro$Edadanos >= 30 & centro$Edadanos <= 39] <- "g. 30-39"
  centro$RangoEtario[centro$Edadanos >= 40 & centro$Edadanos <= 49] <- "h. 40-49"
  centro$RangoEtario[centro$Edadanos >= 50 & centro$Edadanos <= 59] <- "i. 50-59"
  centro$RangoEtario[centro$Edadanos >= 60 & centro$Edadanos <= 69] <- "j. 60-69"
  centro$RangoEtario[centro$Edadanos >= 70] <- "k. > 70"
  centro$RangoEtario[centro$Edadanos == -1] <- "l. Desconocido"

  list(centro = centro)
})

#----#
#Actualizar valores en los filtros de variable
#----#
observe({
  centro <- coledad()$centro

  #UpdateInput: varFilter1
  updateSelectInput(session, "varFilter1",
    choices = sort(unique(centro[, which(colnames(centro)==input$var1)])),
    selected = character(0))
})
observe({
  centro <- coledad()$centro
  #UpdateInput: varFilter2
  updateSelectInput(session, "varFilter2",
    choices = sort(unique(centro[, which(colnames(centro)==input$var2)])),
    selected = character(0))
})
observe({
  centro <- coledad()$centro
  #UpdateInput: varFilter3
  updateSelectInput(session, "varFilter3",
    choices = sort(unique(centro[, which(colnames(centro)==input$var3)])),

```

```

        selected = character(0))
})

#----#
#Filtro tablas Poisson
#----#
centroPoisson <- reactive ({

  centro <- coledad()$centro

  #Evaluar filtro de: varFilter1
  if(is.null(input$varFilter1)){
    fillres <- as.character(unique(centro[, which(colnames(centro)==input$var1)]))
    centro <- subset(centro, centro[,which(colnames(centro)==input$var1)] %in% fillres)
  } else {
    fillres <- input$varFilter1
    centro <- subset(centro, centro[,which(colnames(centro)==input$var1)] %in% fillres)
  }

  #Evaluar filtro de: varFilter2
  if(input$var2 != "-" & !is.null(input$varFilter2)){
    fil2res <- input$varFilter2
    centro <- subset(centro, centro[,which(colnames(centro)==input$var2)] %in% fil2res)
  }

  #Evaluar filtro de: varFilter3
  if(input$var3 != "-" & !is.null(input$varFilter3)){
    fil3res <- input$varFilter3
    centro <- subset(centro, centro[,which(colnames(centro)==input$var3)] %in% fil3res)
  }

  list(centro = centro)
})

#----#
#Reseteo de filtros Poisson
#----#
observeEvent(input$resetPoi,{
  centro <- coledad()$centro
  centro2 <- coledad()$centro

  #Reseteo filtro: varFilter1, usando update para la variable seleccionada
  centro <- centro2
  updateSelectInput(session, "varFilter1",
    choices = sort(unique(centro2[, which(colnames(centro2)==input$var1)])),
    selected = character(0))

  #Reseteo filtro: varFilter2, usando update para la variable seleccionada
  centro <- centro2
  updateSelectInput(session, "varFilter2",
    choices = sort(unique(centro2[, which(colnames(centro2)==input$var2)])),
    selected = character(0))

  #Reseteo filtro: varFilter3, usando update para la variable seleccionada
  centro <- centro2
  updateSelectInput(session, "varFilter3",
    choices = sort(unique(centro2[, which(colnames(centro2)==input$var3)])),
    selected = character(0))
})

#----#
#Definicion de columnas con valor: OTROS
#----#
colsRef <- reactive({
  centro <- centroPoisson()$centro

  var1 <- input$var1
  var2 <- input$var2
  var3 <- input$var3

  #Preparacion para poder definir variables de referencia (ref1, 2 y 3)
  #Sexo y NueveOnce tendran sus valores pre-definidos
  #Provincia, RangoEtario y Severidad usaran el valor maximo
  #Canton, Toxico1, Codtox1, Causa, Ruta y Entidades usaran la variable Otros si se puede
  #calcular con el valor que el usuario provee en num1, 2 y 3.
  #var1
  if(var1 != "Sexo" | var1 == "NueveOnce"){
    colOtrosVar1 <- aggregate(
      centro$NoConsulta,
      by = list(var1 = (centro[, which(colnames(centro)==var1)])),
      FUN=length)
  }

  valores1 <- c()
  centro$ColRef1 <- centro[,which(colnames(centro)==var1)]

```

```

if(var1 == "Provincia" | var1 == "RangoEtario" | var1 == "Severidad"){
  valores1 <- col0trosVar1$var1[col0trosVar1$x == max(col0trosVar1$x)][1]
} else if (var1 == "Canton" | var1 == "Toxico1" | var1 == "Codtox1" | var1 == "Causa" | var1 == "Ruta" | var1 ==
"Entidades") {
  for(i in seq(length(col0trosVar1$var1))){
    if(col0trosVar1$x[i] <= input$num1){
      valores1 <- c(valores1, col0trosVar1$var1[i])
    }
  }
  centro$ColRef1[centro[,which(colnames(centro)==var1)] %in% valores1] <- "Otros"
}
###TODO arreglar el crash que sucede cuando input$num1 is null, no deberia de serlo
#pero el app crashea si se eliminan los numeros y no se ponen otros rapido o si
#dan enter sin culpa antes de poner algun numero.

#var2
if(var2 != "-" & var2 != "Sexo" | var2 == "NueveOnce"){
  col0trosVar2 <- aggregate(
    centro$NoConsulta,
    by = list(var2 = (centro[, which(colnames(centro)==var2)])),
    FUN=length)
}

valores2 <- c()
centro$ColRef2 <- centro[,which(colnames(centro)==var2)]
if(var2 != "-" & var2 == "Provincia" | var2 == "RangoEtario" | var2 == "Severidad"){
  valores2 <- col0trosVar2$var2[col0trosVar2$x == max(col0trosVar2$x)][1]
} else if (var2 != "-" & var2 == "Canton" | var2 == "Toxico1" | var2 == "Codtox1" | var2 == "Causa" | var2 == "Ruta" | var2
== "Entidades") {
  for(i in seq(length(col0trosVar2$var2))){
    if(col0trosVar2$x[i] <= input$num2){
      valores2 <- c(valores2, col0trosVar2$var2[i])
    }
  }
  centro$ColRef2[centro[,which(colnames(centro)==var2)] %in% valores2] <- "Otros"
}
###TODO arreglar el crash que sucede cuando input$num2 is null, no deberia de serlo
#pero el app crashea si se eliminan los numeros y no se ponen otros rapido o si
#dan enter sin culpa antes de poner algun numero.

#var3
if(var3 != "-" & var3 != "Sexo" | var3 == "NueveOnce"){
  col0trosVar3 <- aggregate(
    centro$NoConsulta,
    by = list(var3 = (centro[, which(colnames(centro)==var3)])),
    FUN=length)
}

valores3 <- c()
centro$ColRef3 <- centro[,which(colnames(centro)==var3)]
if(var3 != "-" & var3 == "Provincia" | var3 == "RangoEtario" | var3 == "Severidad"){
  valores3 <- col0trosVar3$var3[col0trosVar3$x == max(col0trosVar3$x)][1]
} else if (var3 != "-" & var3 == "Canton" | var3 == "Toxico1" | var3 == "Codtox1" | var3 == "Causa" | var3 == "Ruta" | var3
== "Entidades") {
  for(i in seq(length(col0trosVar3$var3))){
    if(col0trosVar3$x[i] <= input$num3){
      valores3 <- c(valores3, col0trosVar3$var3[i])
    }
  }
  centro$ColRef3[centro[,which(colnames(centro)==var3)] %in% valores3] <- "Otros"
}
###TODO arreglar el crash que sucede cuando input$num3 is null, no deberia de serlo
#pero el app crashea si se eliminan los numeros y no se ponen otros rapido o si
#dan enter sin culpa antes de poner algun numero.

list(centro = centro, valores1 = valores1, valores2 = valores2, valores3 = valores3)
})

#----#
#Definicion de valores de referencia
#----#
observe({
  centro2 <- colsRef()$centro
  valores1 <- colsRef()$valores1
  valores2 <- colsRef()$valores2
  valores3 <- colsRef()$valores3

#UpdateInput: ref1
var1 <- input$var1
if(var1 == "NueveOnce"){
  updateSelectInput(session, "ref1",
                    #label = "Valor de referencia 1:",
                    choices = sort(unique(centro2$ColRef1)),
                    selected = "No")
} else if(var1 == "Sexo"){

```

```

updateSelectInput(session, "ref1",
                  #label = "Valor de referencia 1:",
                  choices = sort(unique(centro2$ColRef1)),
                  selected = "02.Femenino")
} else if (length(valores1)>1){
updateSelectInput(session, "ref1",
                  #label = "Valor de referencia 1:",
                  choices = sort(unique(centro2$ColRef1)),
                  selected = "Otros")
} else if (length(valores1)==1){
updateSelectInput(session, "ref1",
                  #label = "Valor de referencia 1:",
                  choices = sort(unique(centro2$ColRef1)),
                  selected = valores1)
} else {
updateSelectInput(session, "ref1",
                  #label = "Valor de referencia 1:",
                  choices = sort(unique(centro2$ColRef1)),
                  selected = character(0))
}

#UpdateInput: ref2
if(input$var2 != "-"){
var2 <- input$var2
if(var2 == "NueveOnce"){
updateSelectInput(session, "ref2",
                  #label = "Valor de referencia 2:",
                  choices = sort(unique(centro2$ColRef2)),
                  selected = "No")
} else if(var2 == "Sexo"){
updateSelectInput(session, "ref2",
                  #label = "Valor de referencia 2:",
                  choices = sort(unique(centro2$ColRef2)),
                  selected = "02.Femenino")
} else if (length(valores2)>1){
updateSelectInput(session, "ref2",
                  #label = "Valor de referencia 2:",
                  choices = sort(unique(centro2$ColRef2)),
                  selected = "Otros")
} else if (length(valores2)==1){
updateSelectInput(session, "ref2",
                  #label = "Valor de referencia 2:",
                  choices = sort(unique(centro2$ColRef2)),
                  selected = valores2)
} else {
updateSelectInput(session, "ref2",
                  #label = "Valor de referencia 2:",
                  choices = sort(unique(centro2$ColRef2)),
                  selected = character(0))
}
} else {
updateSelectInput(session, "ref2",
                  #label = "Valor de referencia 2:",
                  choices = "-",
                  selected = "-")
}

#UpdateInput: ref3
if(input$var3 != "-"){
var3 <- input$var3
if(var3 == "NueveOnce"){
updateSelectInput(session, "ref3",
                  #label = "Valor de referencia 3:",
                  choices = sort(unique(centro2$ColRef3)),
                  selected = "No")
} else if(var3 == "Sexo"){
updateSelectInput(session, "ref3",
                  #label = "Valor de referencia 3:",
                  choices = sort(unique(centro2$ColRef3)),
                  selected = "02.Femenino")
} else if (length(valores3)>1){
updateSelectInput(session, "ref3",
                  #label = "Valor de referencia 3:",
                  choices = sort(unique(centro2$ColRef3)),
                  selected = "Otros")
} else if (length(valores3)==1){
updateSelectInput(session, "ref3",
                  #label = "Valor de referencia 3:",
                  choices = sort(unique(centro2$ColRef3)),
                  selected = valores3)
} else {
updateSelectInput(session, "ref3",
                  #label = "Valor de referencia 3:",
                  choices = sort(unique(centro2$ColRef3)),
                  selected = character(0))
}
}

```

```

    }
  } else {
    updateSelectInput(session, "ref3",
                      #label = "Valor de referencia 3:",
                      choices = "-",
                      selected = "-")
  }
})

#----#
#Conteo de intoxicaciones segun las variables elegidas
#----#
PrepVariables <- reactive({
  centro2 <- colsRef()$centro

  #total de intoxicaciones por variables seleccionadas
  if(input$var2 != "-" & input$var3 != "-"){
    intoxicacionPorVariables <- aggregate(
      centro2$NoConsulta,
      by=list(var1 = centro2$ColRef1,
             var2 = centro2$ColRef2,
             var3 = centro2$ColRef3),
      FUN=length)

    #Definir el nivel de referencia
    intoxicacionPorVariables$var1 <- relevel(as.factor(intoxicacionPorVariables$var1), ref = input$ref1)
    intoxicacionPorVariables$var2 <- relevel(as.factor(intoxicacionPorVariables$var2), ref = input$ref2)
    intoxicacionPorVariables$var3 <- relevel(as.factor(intoxicacionPorVariables$var3), ref = input$ref3)
  } else if(input$var2 != "-" & input$var3 == "-") {
    intoxicacionPorVariables <- aggregate(
      centro2$NoConsulta,
      by=list(var1 = centro2$ColRef1,
             var2 = centro2$ColRef2),
      FUN=length)

    #Definir el nivel de referencia
    intoxicacionPorVariables$var1 <- relevel(as.factor(intoxicacionPorVariables$var1), ref = input$ref1)
    intoxicacionPorVariables$var2 <- relevel(as.factor(intoxicacionPorVariables$var2), ref = input$ref2)
  } else if (input$var2 == "-" & input$var3 != "-") {
    intoxicacionPorVariables <- aggregate(
      centro2$NoConsulta,
      by=list(var1 = centro2$ColRef1,
             var3 = centro2$ColRef3),
      FUN=length)

    #Definir el nivel de referencia
    intoxicacionPorVariables$var1 <- relevel(as.factor(intoxicacionPorVariables$var1), ref = input$ref1)
    intoxicacionPorVariables$var3 <- relevel(as.factor(intoxicacionPorVariables$var3), ref = input$ref3)
  } else {
    intoxicacionPorVariables <- aggregate(
      centro2$NoConsulta,
      by=list(var1 = centro2$ColRef1),
      FUN=length)

    #Definir el nivel de referencia
    intoxicacionPorVariables$var1 <- relevel(as.factor(intoxicacionPorVariables$var1), ref = input$ref1)
  }

  #Renombramos las columna X
  colnames(intoxicacionPorVariables)[colnames(intoxicacionPorVariables)=="x"] <- "NumIntoxicaciones"

  list(intoxicacionPorVariables = intoxicacionPorVariables)
})

#----#
#Definicion de Poisson (glm), definicion de IRR (usando deltamethod del paquete msm)
#----#
calcPoisson <- reactive({
  numIntox <- PrepVariables()$intoxicacionPorVariables
  # numVariables <- PrepVariables()$numVariables

  #Definicion de la formula para Poisson con o sin interacciones (glm)
  if(input$intPoisson == "No"){
    #Sin interacciones
    if(input$var2 != "-" & input$var3 != "-"){
      p1 <- glm(NumIntoxicaciones ~ var1 + var2 + var3, family=poisson, data = numIntox)
    } else if(input$var2 != "-" & input$var3 == "-") {
      p1 <- glm(NumIntoxicaciones ~ var1 + var2, family=poisson, data = numIntox)
    } else if (input$var2 == "-" & input$var3 != "-") {
      p1 <- glm(NumIntoxicaciones ~ var1 + var3, family=poisson, data = numIntox)
    } else {

```

```

    p1 <- glm(NumIntoxicaciones ~ var1, family=poisson, data = numIntox)
  }
} else {
#Con interacciones
if(input$var2 != "-" & input$var3 != "-"){
  p1 <- glm(NumIntoxicaciones ~ var1 * var2 * var3, family=poisson, data = numIntox)
} else if(input$var2 != "-" & input$var3 == "-") {
  p1 <- glm(NumIntoxicaciones ~ var1 * var2, family=poisson, data = numIntox)
} else if (input$var2 == "-" & input$var3 != "-") {
  p1 <- glm(NumIntoxicaciones ~ var1 * var3, family=poisson, data = numIntox)
} else {
  p1 <- glm(NumIntoxicaciones ~ var1, family=poisson, data = numIntox)
}
}
}

#Prueba de sobredispersión
#Función Lineal
sobreLin <- dispersiontest(p1, trafo = 1)
#Función Cuadrática
sobreCua <- dispersiontest(p1, trafo = 2)

#Tabla de sobredispersiones
sobreTab <- rbind(sobreLin, sobreCua)
sobreTab <- sobreTab[,c(3,1,2)]
sobreTab <- apply(sobreTab, 2, as.numeric)
sobreTab <- round(sobreTab, 3)
colnames(sobreTab) <- c("Dispersion", "z", "Valor p")
rownames(sobreTab) <- c("Func. Lineal", "Func. Cuadrática")

#Definición de IRR (usando deltamethod del paquete msm)
#Número de variables utilizado = numVariables = # de líneas de coeficientes

#Cálculo de errores robustos y los límites de confianza
cov.p1 <- vcovHC(p1, type="HC0")
std.err <- sqrt(diag(cov.p1))
coef <- na.omit(coef(p1))
r.est <- cbind(IRR= exp(coef),
              "Errores Robustos" = exp(coef)*std.err,
              LI = exp(coef - 1.96 * std.err),
              LS = exp(coef + 1.96 * std.err),
              "Valor p" = 2 * pnorm(abs(coef/std.err), lower.tail=FALSE))

rexp.est <- as.data.frame(round(r.est,3))
rexp.est$Significativo <- "No"
rexp.est$Significativo[as.numeric(rexp.est[,5]) < 0.05] <- "Si"

#Cálculo de la Prueba de bondad de ajuste
# bondAjust <- with(p1, cbind(res.deviance = deviance, df = df.residual,
#                             p = pchisq(deviance, df.residual, lower.tail=FALSE)))

list(sobreTab = sobreTab, p1 = p1, rexp.est = rexp.est) #, bondAjust = bondAjust
})

#Valor de AIC
output$AIC <- renderValueBox({
  p1 <- calcPoisson()$p1

  valueBox(
    round(summary(p1)$aic, 3),
    "AIC",
    icon = icon("chart-scatter"),
    color = "navy"
  )
})

#Valor de R2 de McFadden
#TODO calcular el R2 y actualizar
output$R2 <- renderValueBox({
  p1 <- calcPoisson()$p1

  valueBox(
    round(RsqGLM(p1)$McFadden, 3),
    "R2 (McFadden)",
    icon = icon("chart-scatter"),
    color = "teal"
  )
})

#Render de prueba de sobredispersión
output$sobredisp <- renderDataTable({
  sobreTab <- calcPoisson()$sobreTab

  datatable(round(sobreTab, 3), options = list(scrollX = TRUE))
})

```

```

#Render de la tabla para bondad de ajuste
# output$bondAjust <- renderDataTable({
#   bondAjust <- calcPoisson()$bondAjust
#   colnames(bondAjust) <- c("Devianza Residual", "GL", "Valor p")
#
#   datatable(round(bondAjust,3), options = list(scrollX = TRUE))
# })

#Render de la tabla de Factor de Inflacion de la varianza
output$factInfl <- renderDataTable({
  p1 <- calcPoisson()$p1
  if(!is.null(input$var1) & "-" != input$var2 | !is.null(input$var1) & "-" != input$var3 |
    "-" != input$var2 & "-" != input$var3){
    factInflVar <- vif(p1)
    factInflVar <- cbind(factInflVar, col = factInflVar[,3]^2)
    colnames(factInflVar) <- c("GFIV", "GL", "GFIV^(1/2GL)", "GFIV^(1/GL)")

    datatable(round(factInflVar, 3), options = list(scrollX = TRUE))
  } else {
    factInflVar <- cbind("Se requieren la seleccion de otra variable para calcular el FIV")
    colnames(factInflVar) <- "Restriccion"
    datatable(factInflVar, options = list(scrollX = TRUE))
  }
})

#Render tabla de Incident Rate Ratio
output$IRR <- renderDataTable({
  rexp.est <- calcPoisson()$rexp.est

  datatable(rexp.est, options = list(scrollX = TRUE, pageLength = 20))
})

#Render resumen de detalles del Poisson
output$detallesPoi <- renderPrint({
  p1 <- calcPoisson()$p1

  summary(p1)
})
}

#-----
#Shiny App
#-----
shinyApp(ui = ui, server = server)

```