



MODELOS Y METODOLOGÍAS

PARA LA SOLUCIÓN DE PROBLEMAS
DE IMPACTO SOCIAL Y PRODUCTIVO

*En memoria del **Dr Alfonso Aldape Alamillo**,
decano del TecNM-IT Cd Juárez, México.*

Coordinadores:

Manuel Arnoldo Rodríguez Medina
Iván Juan Carlos Pérez Olguín
Eduardo Rafael Poblano Ojinaga



Modelos y metodologías para la solución de problemas de impacto social y productivo

ISBN (CENID): 978-607-8830-13-8

ISBN (AEVA): 978-84-09-46635-1

Primera edición, 2022

Todos los derechos reservados.

© 2022, Manuel Arnoldo Rodríguez Medina, Iván Juan Carlos Pérez Olguin, Eduardo Rafael Poblano Ojinaga, coordinadores.

© 2022, Manuel Arnoldo Rodríguez Medina, Alejandra Flores Sánchez, Eduardo Rafael Poblano Ojinaga, Francisco Zorrilla Briones, Yuliana Meléndez Pastrana, Alfonso Aldape Alamillo,[†] Juan Pedro Benítez Guadarrama, Lizette Alvarado Tarango, Jesús Elisabeth Sánchez Padilla, Adrian Francisco Loera Castro, Brenda Pedroza Figueroa, Jorge Adolfo Pinto Santos, Miguel Ángel Hernández Rivera, Genoveva Cruz Hernández, Joeovany Rafael Rodríguez Mejía, Manuel Iván Rodríguez Borbón, Perla Ivette Gómez Zepeda, autores.

Los conceptos expresados en este documento son responsabilidad exclusiva de los autores. Esta obra cumple con el requisito de evaluación por dos pares de expertos.

Edición: Julio Rivas Rojas.

Diagramación: Felipe Ponce.

Editorial Centro de Estudios e Investigaciones para el Desarrollo Docente. CENID AC es miembro de la Cámara Nacional de la Industria Editorial Mexicana Socio #3758.

Queda prohibida la reproducción o transmisión total o parcial del contenido de la presente obra mediante algún método sea electrónico o mecánico (INCLUYENDO EL FOTOCOPIADO, la grabación o cualquier sistema de recuperación o almacenamiento de información), sin el consentimiento por escrito del editor.

Indexación de datos

Bases de datos en las que *Editorial Centro de Estudios e Investigaciones para el Desarrollo Docente CENID A.C.* está indexada: Dialnet (Universidad de la Rioja).

© 2022 Editorial Centro de Estudios e Investigaciones para el Desarrollo Docente. CENID AC Pompeya # 2705. Colonia Providencia C.P. 44670 Guadalajara, Jalisco. México Teléfono: 01 (33) 1061 8187 Registro Definitivo Reniecyt No.1700205 a cargo de Conacyt.

© 2022 Editorial de la Asociación Científica para la Evaluación y Medición del los Valores Humanos c/ de les cases sert nº 11, C.P. 08193, Bellaterra – Cerdanyola del Vallés (Barcelona).

CENID y su símbolo identificador son una marca comercial registrada.

Impreso en México / Printed in México



Si desea publicar un libro o un artículo de investigación contáctenos.

www.cenid.org

redesdeproduccioncenid@cenid.org

| | |
|--|------------|
| Semblanza del Dr. Alfonso Aldape Alamillo | 5 |
| Introducción | 7 |
| Construcción de modelo para determinar factores significantes en la adquisición del Virus del Papiloma Humano | 10 |
| Manuel Arnoldo Rodríguez Medina, Alejandra Flores Sánchez, Eduardo Rafael Poblano Ojinaga | |
| Contraste metodológico entre un algoritmo de búsqueda directa y superficie de respuesta | 33 |
| Francisco Zorrilla Briones, Yuliana Meléndez Pastrana, Alfonso Aldape Alamillo | |
| Desarrollo Organizacional: Un Modelo | 63 |
| Alfonso Aldape Alamillo,† Francisco Zorrilla Briones, Juan Pedro Benítez Guadarrama | |
| Desnutrición y apoyo a comunidades marginadas y el impacto de su programa de RSE: caso de una empresa del norte de México | 80 |
| Lizette Alvarado Tarango, Jesús Elisabeth Sánchez Padilla, Adrian Francisco Loera Castro | |
| Evaluación de la inteligencia competitiva y la capacidad de innovación en empresas localizadas en la Región Lagunera-México | 106 |
| Eduardo Rafael Poblano Ojinaga, Brenda Pedroza Figueroa, Jorge Adolfo Pinto Santos | |
| Identificación de los factores críticos para mejorar la eficacia de la cadena de suministro: Una propuesta | 132 |
| Miguel Ángel Hernández Rivera, Genoveva Cruz Hernández, Manuel Arnoldo Rodríguez Medina | |
| Modelado estadístico en pruebas aceleradas de confiabilidad en mantenimiento | 151 |
| Jorge Adolfo Pinto Santos, Joeovany Rafael Rodríguez Mejía, Manuel Iván Rodríguez Borbón | |

**Resolución de un plan de distribución a través del razonamiento
basado en casos y redes bayesianas**

Adrián Francisco Loera Castro, Perla Ivette Gómez Zepeda,
Lizette Alvarado Tarango

174

Construcción de modelo para determinar factores significantes en la adquisición del virus del papiloma humano (VPH)

Construction of a model to determine significant factors in the acquisition of the Human Papillomavirus (HPV)

Manuel Arnoldo Rodríguez Medina

Tecnológico Nacional de México / IT de Ciudad Juárez, México

manuel.rm1@cdjuarez.tecnm.mx

<https://orcid.org/0000-0002-8922-4718>

Alejandra Flores Sánchez

Universidad Autónoma de Ciudad Juárez, México

alejandra.flores@uacj.mx

<https://orcid.org/0000-0002-2002-1330>

Eduardo Rafael Poblano Ojinaga

Tecnológico Nacional de México / IT de Ciudad Juárez, México

eduardo.po@cdjuarez.tecnm.mx

<https://orcid.org/0000-0003-3482-7252>

Resumen

Este documento explora el conocimiento sobre cómo desarrollar un modelo para determinar la confiabilidad del diagnóstico del virus del papiloma humano (VPH) mediante inferencia bayesiana con el fin de revelar los factores significativos que vuelven vulnerables a la población femenina (objeto de estudio en la investigación). Lo anterior debido a que se han desarrollado diversos estudios para determinar cómo erradicarlo, pero no el cómo evitar el problema o eliminarlo desde la raíz, ya que no se tienen identificadas con exactitud las variables que originan dicho virus. El primer objetivo del proyecto fue la obtención de la información, de tal manera que fuese suficientemente eficaz y confiable. Esto se logró mediante una campaña de concientización en mujeres. La información requerida se obtuvo aplicando una encuesta a las pacientes que acudieron al centro de análisis sobre la presencia del VPH. Luego se clasificaron como saludables las que no tenían el virus y no saludables las que lo habían adquirido. Después se aplicaron herramientas estadísticas pertinentes: primero la descriptiva y posteriormente el teorema de Bayes. La importancia del estudio radica en que se identificaron

aquellas variables que pueden ayudar a prevenir y disminuir la incidencia de VPH en las mujeres, principal causa del cáncer cervicouterino (CaCu) a nivel mundial. Finalmente, se generó un modelo para determinar la confiabilidad del diagnóstico del VPH mediante inferencia bayesiana, clave para demostrar que con inferencia bayesiana es posible predecir con base en evidencia existente. Probablemente, lo más convincente debe ser que ofrece una manera atractiva e intuitiva, pero metodológicamente rigurosa, de actualizar nuestro conocimiento sobre el tema. Además, es importante aclarar que no es lo sofisticado de algunos equipos lo que justifica la presencia de profesionales de las ciencias exactas en la medicina, sino la diversidad de tecnologías, su utilización, los estudios de costo-beneficio, y la confiabilidad del diagnóstico y tratamiento en las diversas enfermedades que aquejan al ser humano, esto con el fin de lograr el bienestar del hombre. La investigación sustenta, entre otras implicaciones, la necesidad de aplicar inferencia bayesiana para desarrollar, implementar y evaluar otras áreas.

Palabras clave: estadística bayesiana, virus del papiloma humano, confiabilidad, salud pública.

Abstract

This document explores the knowledge on how to develop a model to determine the reliability of the diagnosis of the Human Papilloma Virus (HPV) through Bayesian inference, an inference that reveals the significant factors that make the female population vulnerable, which were the object of study in the research; This is due to the fact that several studies have been carried out to determine how to eradicate it, but not how to avoid the problem or eliminate it from the root, since the variables that cause said virus have not been exactly identified. The first objective of the project was to obtain the information, in such a way that it would be sufficiently effective and reliable. This was achieved through an awareness campaign in women, considering that it is the part of the population most affected by HPV. The required information was obtained by applying a survey to the patients who attended the analysis center on the presence of HPV. Immediately, those who did not have the virus were classified as healthy and those who had acquired it were classified as unhealthy. After the classification, relevant statistical tools were applied, starting with the descriptive one, and later the application of Bayes' Theorem, and, of course, the use of Bayesian inference. The importance of the study lies in the identification of those variables that can help prevent and reduce the incidence of HPV in women, since it is the main cause of Cervical Cancer (CACU)

worldwide. Finally, a model was generated to determine the reliability of HPV diagnosis through Bayesian inference, key to demonstrate that with Bayesian inference it is possible to predict based on existing evidence. Probably most convincing is that it offers an attractive and intuitive, yet methodologically rigorous way to update our knowledge on the subject. In addition, it is important to clarify that it is not the sophistication of some equipment that justifies the presence of exact science professionals in medicine, but rather the diversity of technologies, their use, cost-benefit studies, and the reliability of diagnosis and treatment. treatment in the various diseases that afflict the human being, this to achieve the well-being of man. The research supports, among other implications, the need to apply Bayesian inference to develop, implement and evaluate other areas.

Keywords: *Bayesian Statistics, Human Papilloma Virus, Reliability, Public Health.*

1. Introducción

En esta sección se explicarán las condiciones en las que se encuentra la mujer en Ciudad Juárez en lo que respecta al virus del papiloma humano (VPH), así como los antecedentes, el planteamiento del problema, el marco teórico, los materiales, los métodos y la bibliográfica revisada hasta el momento con el fin de tener una visión más amplia del caso.

El VPH como causa de cáncer de cuello de útero ocupa el cuarto lugar entre los tipos más comunes de cáncer (a nivel mundial) que afectan a mujeres con un número estimado de 266 000 muertes y unos 528 000 nuevos casos en 2012. La gran mayoría (alrededor de 85 %) de esas muertes se produjeron en las regiones menos desarrolladas, donde se registra casi 12 % de todos los cánceres femeninos.

Aunque muchas de las infecciones por VPH no causan síntomas, la infección genital persistente por VPH puede provocar cáncer de cuello uterino en las mujeres. De hecho, el VPH también puede ocasionar otros tipos de cáncer ano-genital, cáncer de cabeza, cuello y verrugas genitales tanto en hombres como en mujeres. Las infecciones por VPH se transmiten a través del contacto sexual; sin embargo, existen factores asociados que ayudan a que las mujeres sean más susceptibles a contraer el VPH.

2. Planteamiento del problema

2.1 Antecedentes

El VPH es la causa del cáncer de cuello del útero, el cual ocupa el cuarto lugar entre los tipos más comunes de cáncer que afectan a mujeres. Prácticamente todos los casos de esta enfermedad (99 %) están vinculados con la infección genital por el VPH, una de las más comunes del aparato reproductor.

La prevalencia más alta de VPH oncogénico de alto riesgo son los tipos 16, 18, 31, 33, 35, 45, 51, 52, 58, 59, que se encuentran en África y América Latina. El VPH 16 es el más común en el mundo, excepto en Indonesia y Argelia, donde el VPH 18 es el más habitual, mientras que el VPH 45 tiene una alta frecuencia en África Occidental. Los tipos 33, 39 y 59 se concentran en América Central y en América Sur (Rivera *et al.*, 2012).

En México, el cáncer cervicouterino (CaCu) es el más frecuente y cobra 12 víctimas al día. El Centro Nacional de Equidad de Género y Salud Reproductiva (2007) estima que anualmente se diagnostican 22 000 nuevos casos y determina que 50 % de las personas sexualmente activas serán infectadas tarde o temprano por el VPH, lo que conlleva en su gran mayoría a padecer CaCu.

En el caso específico de Ciudad Juárez, confluyen factores que dinamizan la movilidad poblacional, lo que, por una parte, activa la inmigración de connacionales y extranjeros de diferentes regiones nacionales y de otras latitudes del extranjero y, por otra, se produce el fenómeno social de la emigración que genera el movimiento y tráfico de seres humanos hacia Estados Unidos.

Es importante indicar que Ciudad Juárez es la segunda localidad con mayor número de casos de VPH, al detectarse 312 mujeres infectadas de 654 922 —según el Instituto Nacional de Estadística Geografía e Informática [INEGI] (2007)—, mientras que la ciudad de Chihuahua encabeza el listado con un total 649 y Guachochi ocupa el tercer lugar con 92 infectadas.

Flores-Sánchez (2006), en una aplicación de herramientas de ingeniería (Mahalanobis-Taguchi) para identificar los factores de riesgo del VPH, encontró que de 35 variables analizadas en 400 pacientes del Hospital de la Familia en Ciudad Juárez, Chihuahua, en el área de colposcopia, solo resultaron significativas las siguientes: primer periodo menstrual, si el compañero está circuncidado o no; número de embarazos, si la mujer inicia su vida sexual entre los 12 y 13 años, es más propensa a adquirir el virus; si el compañero no está circuncidado, aumenta la probabilidad de adquirirlo, y entre menor sea el número de embarazos mayor es la probabilidad de adquirir el virus.

2.2 Descripción del problema

El problema de investigación considerado es un tema de interés social debido a que el género femenino es el más afectado por este padecimiento; además, los procesos de detección en ocasiones son tardíos y no se le da la importancia requerida.

2.3 Pregunta de investigación

¿Qué factores serían significantes en un sistema modelado mediante análisis bayesiano que explique la evolución del VPH?

2.4 Hipótesis y variables de investigación

H1. Los factores significativos en las pacientes con VPH son edad, peso, estatura, primer periodo menstrual, inicio de vida sexual, compañeros sexuales, compañeros circuncidados, duración de periodo menstrual, embarazos, estatus socioeconómico y vida clínica. Con el análisis bayesiano se puede desarrollar un sistema que modele la extensión del virus.

2.5 Objetivos

1. Desarrollar un modelo para determinar el diagnóstico del VPH mediante inferencia bayesiana.
2. Determinar si la edad, peso, estatura, primer periodo menstrual, inicio de vida sexual, compañeros sexuales, compañeros circuncidados, duración de periodo menstrual, embarazos, estatus socioeconómico y vida clínica son factores de riesgo que determinan un patrón de comportamiento para contraer el VPH y si la inferencia bayesiana es un buen método para la creación de un modelo de comportamiento.

3. Marco teórico

En este apartado se muestra la literatura encontrada hasta el momento referente a diversas metodologías de análisis e investigación del tema de interés.

3.1 VPH y cáncer cervicouterino (CaCu)

Muñoz *et al.* (1993) determinaron que el carcinoma invasor del cuello uterino y sus precursores son causados la mayor parte de las veces por el VPH. Para ello, realizaron cuatro estudios concurrentes de casos y controles: dos de ellos incluyeron casos de

carcinoma invasor y controles de base poblacional, y otros dos incluyeron casos y controles de carcinoma *in situ*. Con este estudio se dieron cuenta de que la exposición al VPH fue el factor de riesgo principal en los cuatro estudios. En la tabla 3.1 se pueden observar los porcentajes de casos y controles positivos al VPH y los riesgos relativos (RR) de la asociación entre VPH y el CaCu en España y Colombia (1985-1988).

Tabla 3.1. Porcentajes de casos y controles positivos al VPH

| | España | | | Colombia | | |
|------------------------------------|-------------------|-----------|--------------|-------------------|-----------|------------|
| | VPH positivos (%) | | RR | VPH positivos (%) | | RR |
| Tipo de cáncer | Casos | Controles | (IC 95%) | Casos | Controles | (IC 95%) |
| Cáncer invasor ^b | 69 | 4.6 | 46.2 | 72.4 | 13.3 | 15.6 |
| | | | (18.5-115.1) | | | (6.9-34.7) |
| Cáncer <i>in situ</i> ^c | 70.7 | 4.7 | 56.9 | 63.2 | 10.5 | 15.5 |
| | | | (24.8-130.6) | | | (8.2-29.4) |

^b Riesgo relativo ajustado según la edad, lugar de estudio, número de compañeros sexuales, edad del primer nacimiento, grado de escolaridad y PAP de tamizaje.

^c Riesgo relativo ajustado según la edad, lugar de estudio, número de compañeros sexuales, edad en el primer coito y número de compañeros sexuales del marido. VPH = virus del papiloma humano.

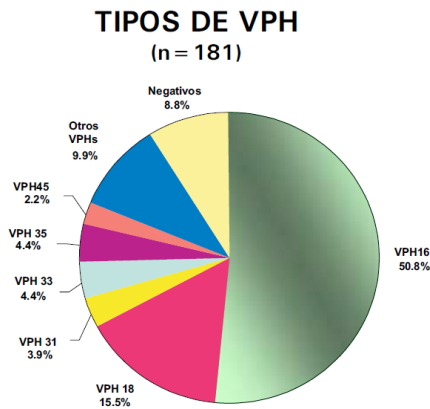
IC 95 % = Intervalo de confianza de 95 %.

Serman (2002) encontró una fuerte asociación entre la infección ano-genital por VPH y el desarrollo de neoplasia intraepitelial cervical y cáncer cervical invasivo. Asimismo, En 2006, frotis de cuello uterino, se descubrió que el ADN del virus del VPH 16 era altamente oncogénico en 47 % de las neoplasias intraepiteliales de alto grado (NIE) y en 47 % de los cánceres de cuello uterino invasivos. En un estudio prospectivo en más de 18 000 mujeres se halló, utilizando técnicas de inmunología molecular, que la presencia de anticuerpos contra el VPH16 confiere un riesgo 12 veces mayor que el resto de la población de desarrollar CaCu o carcinoma *in situ*. Asimismo, se halló que el riesgo es mayor para las neoplasias que se desarrollan 5 años o más después de tomar una muestra de sangre para el estudio de anticuerpos, lo que apoya firmemente el concepto de que la infección persistente por HPV16 (y probablemente cualquier HPV oncogénico) está incidentalmente implicada en la etiología del cáncer de cuello uterino.

Berumen-Campos (2003) describe la alta frecuencia del VPH 16 (figura 3.1 *Tipos de VPH*) y sus variantes de virus asiático-americano en México son más oncogénicas que las europeas. El 23 % de los CaCu son ocasionados por el VPH. Lo anterior lo

descubrieron realizando una hibridación de ácidos nucleicos (identificación de bandas de ADN) en muestras de pacientes.

Figura 3.1. Tipos de VPH



Flores *et al.* (2015) mencionan que la infección por VPH es el principal factor asociado al desarrollo de CaCu. Conocer la prevalencia de los distintos VPH en las diferentes etapas del desarrollo de CaCu es relevante para determinar los VPH de riesgo oncogénico, establecer el desarrollo de estrategias de cribado y la evaluación de programas de prevención, así como para el diseño de vacunas.

3.2 Factores del VPH

Zamudio *et al.* (2001) realizaron un estudio donde compararon la eficacia de la citología y colposcopia en el diagnóstico de las lesiones cervicales del VPH; los autores querían determinar la sensibilidad y especificidad del PAP en una clínica de displasias, así mismo de la colposcopia y relacionar ambos hallazgos con el diagnóstico definitivo (histopatológico), así como conocer los factores de riesgo de mayor prevalencia en la población de estudio. Los resultados (tabla 3.12) que se obtuvieron fue que la colposcopia tiene una sensibilidad de 83 % contra el 41 % del PAP y una especificidad de 66 % menor que 86% del frotis. Los factores de riesgo (tabla 3.13) hallados coinciden con los informados en el resto del mundo: número de parejas sexuales, multiparidad e iniciación temprana en la vida sexual.

Tabla 3.12 Sensibilidad y especificidad en los métodos de diagnóstico

| | Colposcopia | Papanicolau |
|----------------------|-------------|-------------|
| Sensibilidad | 83 % | 41 % |
| Especificidad | 66 % | 86 % |

Tabla 3.13 Factores de riesgo

| | Grupo I | Grupo II |
|----------------------------------|----------------|-----------------|
| N.º de parejas sexuales | 41.90 % | 25 % |
| Tabaquismo | 40.30 % | 33.80 % |
| Multiparidad | 37.60 % | 32.30 % |
| IVSA antes de los 19 años | 27.60 % | 22 % |
| Anticonceptivos orales | 22.60 % | 10.50 % |
| Inmunosupresión | 4.20 % | 1.40 % |

Stanczuk *et al.* (2003) realizaron un estudio sobre la posibilidad de detectar el VPH mediante una prueba de orina. Hoy en día se sabe que la mayoría de los cánceres de cuello uterino invasivos están asociados con la presencia crónica del VPH en el cuello uterino, en particular ciertas variedades con propiedades oncogénicas, como los tipos 16 y 18. En los países en desarrollo, sin embargo, sigue siendo un desafío garantizar que las mujeres, especialmente en las zonas rurales, se someten a exámenes pélvicos periódicos para obtener un frotis cervical, a través del cual se obtiene el material necesario para la detección de virus y CaCu.

La falta de acceso a los servicios de salud, la escasez de recursos humanos y de laboratorio, el miedo de las mujeres a los exámenes ginecológicos y el desconocimiento de la utilidad de la prueba son algunos de los factores que explican la poca frecuencia con la que se realizan frotis de cuello uterino en determinadas partes del mundo en desarrollo. Sería ideal encontrar otro medio de detección viral que sea más fácil de aplicar, menos engorroso y que no requiera de personal capacitado ni de instrumentos especiales para obtener la muestra. Gracias a un estudio realizado en Zimbabue por investigadores de la Universidad de Harare, tal prueba pronto se hará realidad.

3.3 Estadística bayesiana en medicina: 25 años de revisión

Ashby *et al.* (1993) examinaron el estado del pensamiento bayesiano cuando se lanzó *Statistics in Medicine* en 1982, lo que refleja particularmente su aplicabilidad y usos en la investigación médica. Luego analizaron cada época subsiguiente de cinco años, con un enfoque en los artículos que aparecen en *Statistics in Medicine*, poniéndolos en el contexto de los principales desarrollos en el pensamiento bayesiano y la computación con referencia a libros importantes, reuniones históricas y diversos documentos. Así, trazaron el crecimiento de las estadísticas bayesianas a medida que se aplica a la medicina y formuló predicciones para el futuro.

A partir de escasos comienzos, donde apenas se mencionaban las estadísticas bayesianas, ahora impregnan todas las áreas principales de las estadísticas médicas, incluidos ensayos clínicos, epidemiología, meta análisis y síntesis de pruebas, modelado espacial, modelado longitudinal, modelos de supervivencia, genética molecular y decisión del qué hacer respecto de las nuevas tecnologías.

Por toda la investigación que se realizó en este artículo, Ashby *et al.* (1993) predicen que las tendencias de la investigación sugieren que Bayes será utilizado cada vez más por aquellos que no se clasifican a sí mismos como estadísticos, así como aquellos que sí lo hacen, ya que será utilizado en áreas más nuevas y de rápido desarrollo, donde se requiere flexibilidad e innovación.

3.8 Modelado estadístico

Para hablar de modelo lineal generalizado debemos explicar antes el modelado estadístico, que se refiere a la estimación de parámetros, así como la comparación y ajuste de modelos de probabilidad de datos empíricos. Este modelo pretende explicar la variación de una respuesta a partir de la relación conjunta de dos fuentes de variabilidad: una de carácter determinista y la otra aleatoria, lo que responde a la ecuación (3.1).

$$\text{Respuesta} = \text{componente sistemático} + \text{componente aleatorio} \quad (3.1)$$

La construcción, formulación y ajuste de modelos a datos empíricos se encarga del modelado estadístico precisamente, debiendo responder a tres criterios:

- a. Principio de bondad de ajuste: La inclusión de parámetros en el *modelo* en beneficio de una mejor representación de *datos* con la correspondiente disminución del *error*.
- b. Principio de parsimonia: La selección de parámetros que formen el modelo de tal modo que este se convierta en la representación simple y sobria de la realidad.
- c. Integración teórica del modelo en la red conceptual que lo generó.

Una construcción de modelo que explique la variable de respuesta con el menor error posible se realiza basándose en las siguientes etapas:

1. Especificación del modelo teórico: Determina qué variables son de interés, así como cuáles son las relaciones entre ellas.
2. Estimación de parámetros: Calcula el valor de los coeficientes del modelo examinado a partir de un conjunto de datos observados.

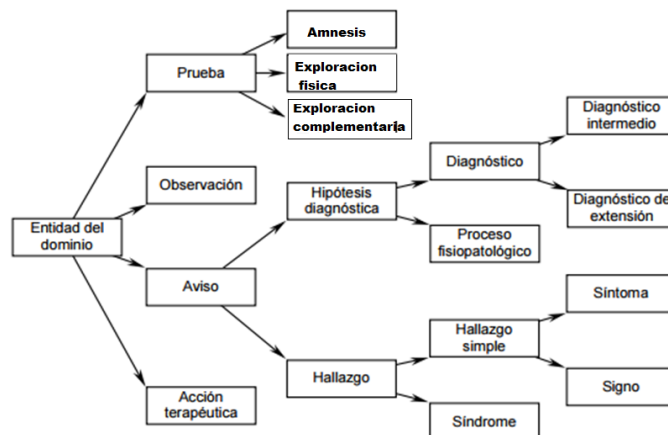
3. Selección del modelo: Valora si el nivel de discrepancia entre lo observado y lo ajustado es suficientemente bajo como para optar por el modelo o rechazarlo.
4. Evaluación del modelo: Se examinan las observaciones individuales, los datos influyentes y los anómalos; asimismo, se comprueban los supuestos de normalidad, linealidad, homocedasticidad e independencia.
5. Interpretación del modelo: Explicación detallada de los parámetros del modelo para comprobar si cumplen los criterios estadístico, lógico y sustantivo.
6. Se acepta o no el modelo, y si es preciso se reinicia el proceso.

4. Metodología

La metodología que se propone desarrollar en esta investigación se muestra en la figura 4.1, en donde se inicia con una entidad del dominio, la cual es la institución hospitalaria en la que se realizó la investigación. Una vez en el área de displacia, se realiza la recolección de historia clínica de la paciente con un objetivo diagnóstico al que se denomina Amnesis, luego se efectúa la colposcopia, que sería la exploración física, y posteriormente se hace la exploración complementaria llevando a patología la muestra que se toma en la exploración física.

Se registran las observaciones y se procede al aviso donde puede ocurrir una hipótesis diagnóstica o un hallazgo.

Figura 4.1 Metodología propuesta



4.1 VPH (*Haemophilus ducrevi*)

La inoculación del virus ocurre por microtraumatismo de los epitelios genitales, masculinos o femeninos, durante el coito con una persona infectada.

El virus penetra a las células de la capa germinativa del epitelio, es decir, a la que da origen al resto de las células que componen el espesor del epitelio cervical y comienzan un proceso de replicación viral. El periodo de incubación del virus varía de 8 semanas a 8 meses, durante los cuales se colonizan grandes extensiones del epitelio genital hasta alcanzar un periodo de infección estable que puede permanecer latente hasta 25 años.

4.2 Vigilancia epidemiológica

Acción de observar, recolectar y analizar sistemáticamente información sobre eventos relacionados con la salud, que permite el conocimiento oportuno y objetivo de los principales problemas de salud y sus factores de riesgo asociados. Es importante realizar pruebas de diagnóstico para la detección oportuna del VPH debido a los beneficios que proporcionan.

4.3 PAP

Es el tamizaje o método que permite realizar un diagnóstico citológico de probable lesión cervical, aunque cualquier hallazgo debe ser confirmado mediante un estudio colposcopio y biopsia dirigida.

4.4 Colposcopia

La colposcopia se realiza por medio de un microscopio especial llamado *colposcopio*. A través de él se consigue ver el cuello del útero al final de la vagina de forma muy ampliada.

Durante la prueba la mujer se acuesta boca arriba en una camilla especial con las piernas separadas. Se introduce un espejo para separar las paredes vaginales. El médico observa el cuello del útero y tiñe su superficie con distintos líquidos, como el ácido acético, para resaltar zonas en las que puede haber células anormales. Estas regiones sospechosas se pueden tomar para estudiarlas posteriormente en una biopsia o, en ocasiones, se pueden eliminar en su totalidad.

Es importante explicar a la paciente el resultado preliminar del estudio. En la figura 4.2 se observan algunos diagnósticos que se pueden encontrar y en la tabla 4.1 se muestra la correlación de las clasificaciones según el método de diagnóstico que se utiliza.

Figura 4.2 Resultados en una colposcopia

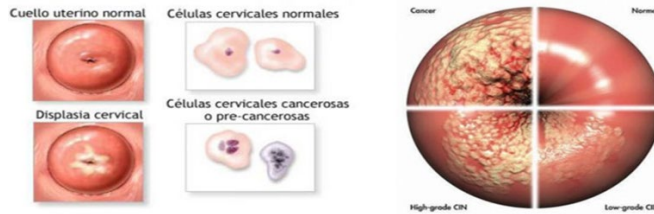


Tabla 4.1 Correlación de las clasificaciones para las lesiones preinvasoras y el cáncer *in situ* del cérvix

vix

| Sistema Bethesda | Normal | | Lesiones intraepiteliales Escamosas de bajo grado | | Lesiones intraepiteliales Escamosas de alto grado | |
|------------------|------------------|-----------------------|---|-------|---|--|
| | Richard | Reagan | Coilocitosis | NIC I | NIC II | NIC III |
| | | | Displasia leve | | Displasia moderada | Displasia severa y cáncer <i>in situ</i> |
| PAP | Clase I Negativo | Clase II Inflamatorio | Clase III Sospechoso | | | Clase IV Altamente sospechoso |

4.5 Seguimiento

Una vez realizado el procedimiento terapéutico, la mujer debe acudir dos semanas después para constatar que la lesión fue removida y que los procesos de cicatrización progresan sin complicación. Posteriormente, se realizan estudios citológicos y colposcópicos cada 3 meses durante el primer año; de no existir ninguna lesión, continuará seguimiento cada 6 meses durante el segundo y tercer año; posteriormente se realizará seguimiento cada año.

4.6 Resultados preliminares

En esta sección se presenta una parte de los resultados obtenidos con las aplicaciones de los métodos a la información.

4.6.1 Análisis por teorema de Bayes

En la actualidad, el teorema de Bayes tiene un gran impacto en la inferencia estadística, así como en sus diversas aplicaciones en muchas ramas de la ciencia, como genética, epidemiología, ingeniería, ciencia forense, medicina, ecología, psicología, entre otras. El teorema de Bayes es ahora una de las piedras fundamentales del trabajo estadístico.

Sean B_1, B_2, \dots, B_k eventos mutuamente excluyentes y exhaustivos. Para cualquier evento A se tiene.

Teorema 1. Teorema de Bayes

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^k P(A|B_i)P(B_i)}$$

Teorema 4.1. Teorema de Bayes para variables aleatorias. Sean X y θ variables aleatorias con fdp $f(x|\theta)$ y $\mathcal{E}(\theta)$

$$\mathcal{E}(\theta|x) = \frac{f(x|\theta)\mathcal{E}(\theta)}{\int_{\theta} f(x|\theta)\mathcal{E}(\theta)d\theta}$$

Dentro del ensayo bayesiano se tiene que:

x : Datos (escalar o vector o matriz)

θ : Parámetro desconocido (escalar o vector o matriz)

$f(x_1, \dots, x_n|\theta)$ verosimilitud de los datos dado el parámetro (desconocido) θ

$\mathcal{E}(\theta)$: Distribución a priori de θ

Por el teorema anterior:

$$\mathcal{E}(\theta|x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n|\theta)\mathcal{E}(\theta)}{\int_{\theta} f(x_1, \dots, x_n|\theta)\mathcal{E}(\theta)d\theta}$$

El aprendizaje bayesiano será:

$$\mathcal{E}(\theta|x_1) \propto f(x_1|\theta) \mathcal{E}(\theta)$$

$$\mathcal{E}(\theta|x_1, x_2) \propto f(x_2|\theta)f(x_1|\theta) \mathcal{E}(\theta) \propto f(x_2|\theta) \mathcal{E}(\theta|x_1)$$

Por lo tanto, el teorema de Bayes muestra cómo el conocimiento acerca del estado de la naturaleza representada por θ es continuamente modificada a medida que nuevos datos son adquiridos (Correa y Barrera, 2018).

Con la aplicación del teorema de Bayes, el conjunto inicial serán los pacientes que acuden a un hospital de Ciudad Juárez y estos se clasificaran en dos tipos: padecer el virus VPH y no padecerlo (NO VPH). Está claro que no tienen elementos comunes (no puede haber nadie que padezca y no padezca la enfermedad a la vez) y que si conjuntamos todos los que padecen dicho virus con los que no lo padecen, obtenemos el conjunto de la población.

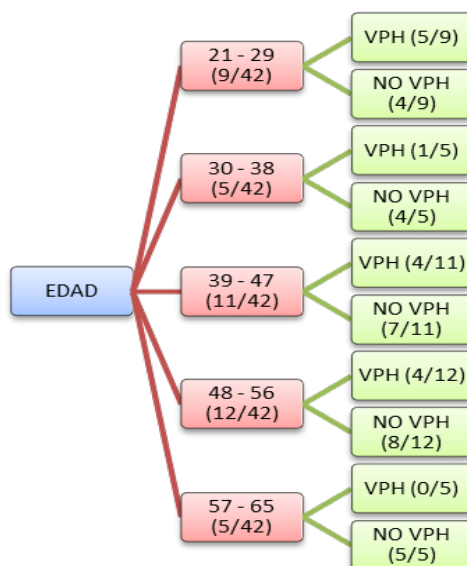
Enseguida, basándose en una entrevista (de 17 preguntas) a las pacientes antes de realizar la prueba de colposcopia, se estableció, *a priori*, que 67 % de las pacientes estaban sanas y 33 % eran portadoras del virus.

A continuación, se muestra un ejemplo de los resultados por variable, el árbol de probabilidades y los cálculos con el teorema de Bayes que se aplicó (tabla 4.2 y figura 4.3).

Tabla 4.2 Frecuencia de acuerdo con la edad

| EDAD | NO VPH | VPH | TOTAL |
|---------|--------|-----|-------|
| 21 – 29 | 4 | 5 | 9 |
| 30 – 38 | 4 | 1 | 5 |
| 39 – 47 | 7 | 4 | 11 |
| 48 – 56 | 8 | 4 | 12 |
| 57 – 65 | 5 | 0 | 5 |
| | 28 | 14 | 42 |

Figura 4.3 Árbol de probabilidad de la variable edad



4.6.2 Hipótesis, verosimilitud y probabilidad *a priori*

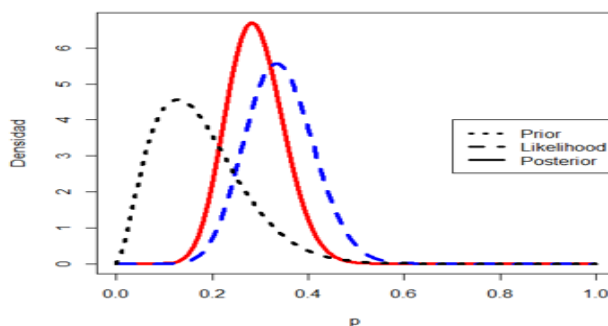
Verosimilitud y probabilidad a posteriori. La fórmula de Bayes es una herramienta matemática que permite utilizar los datos para actualizar las probabilidades del modelo de forma rigurosa y sistemática. La verosimilitud asociada a un modelo es la probabilidad de que sucedan los datos observados bajo las hipótesis del modelo.

La fórmula de Bayes combina la probabilidad *a priori* con la verosimilitud para establecer la probabilidad *a posteriori*. Esta probabilidad solo está disponible después de que los datos hayan sido observados. Lo anterior lo podemos observar en la figura 4.4 y 4.5 donde se muestran los comandos y el gráfico de ellos en el programa R® que se utilizó para el análisis.

Figura 4.4 Programa en R® para determinar verosimilitud *a priori* y *a posteriori*

| | | |
|---|---|--|
| <pre> EDAD en R library(LearnBayes) library(RcmdrMisc) #Estimación proporciones #obtener la prior a partir de dos cuantiles #(50pcenti,90pcentil en #el ejemplo. Se puede #cambiar) la función beta.select da como #resultado #los valores de a y b de la d.beta (d.a priori) quantile1=list(p=.5,x=0.40) quantile2=list(p=.9,x=0.89) param_beta<-beta.select(quantile1,quan- tile2) a=param_beta[1] b=param_beta[2] parámetros.d.beta<-list(param_beta) parámetros.d.beta #dibujo de la distribución a priori .x <- seq(0.023, 0.66, length.out=1000) </pre> | <pre> plotDistr(.x, dbeta(.x, shape1=a, shape2=b), cdf=FALSE, xlab="p", ylab="Densidad") main=paste("D. a priori beta") remove(.x) #muestra exitos<-8 fallos<-15 #plot_verosimilitud .x <- seq(0.023, 0.66, length.out=1000) plotDistr(.x,dbinom(exitos,size= exitos+ fallos,prob=.x),xlab="valor de p",ylab="verosimilitud",main=paste ("verosimilitud")) remove(.x) #Posterior aa=a+ exitos bb=b+ fallos ppost<-c(aa,bb) ppost posterior<-list("beta",ppost) posterior #estimacion supuesta pérdida cuadrática estima<-aa/(aa+bb) estimacion<- </pre> | <pre> print(c("estimacion supuesta pérdida cuadrática",estima)) #dibujo la d.posterior .x <- seq(0.023, 0.66, length.out=1000) plotDistr(.x, dbeta(.x, shape1=aa, shape2=bb), cdf=FALSE, xlab="p", ylab="Density", main=paste(c("D. a posteri- ori", posterior))) remove(.x) #triplot: Plot de prior, veros. y posterior prior=c(a,b) # proporción tiene una prior beta(a, b) data=c(exitos,fallos) # se observan exitos y fa- llos triplot(prior,data) </pre> |
|---|---|--|

Figura 4.5 Grafica de densidad *a priori*, verosimilitud *a posteriori*



4.7 Caso de estudio

De acuerdo con la estimación de la proporción que se ha realizado por medio del teorema de Bayes, se estima que 33 % de las mujeres tamizadas (las que se realizan la prueba del VPH) serán positivas al VPH. Con esta información realizamos la estimación de los cuantiles para determinar los parámetros de la distribución beta.

Se puede calcular la proporción de que las pacientes no saludables sea mayor a 0.5. La figura 4.6 muestra cómo realizarlo en R®:

Figura 4.6 Cálculo de la proporción mayor a 0.5

La probabilidad $P = (p > 0.5 | \text{datos})$
 $1 - \text{pbeta}(0.5, a+s, b+f)$
 0.0005133649

4.7.1 Aplicación de regresión logística

Como aplicación de regresión logística en el entorno bayesiano, analizamos los factores que afectan a pacientes con VPH. La variable de resultado es si el paciente es saludable o no saludable, codificado como 0 o 1, y las variables de edad, peso, estatura, primer periodo menstrual, inicio de vida sexual, número de compañeros sexuales, edad del compañero sexual, días de sangrado, número de embarazos, cesáreas, partos, abortos, número de horas que duerme, días de la semana que consume maíz en sus diferentes presentaciones, número de días que consume carnes rojas, número de días que consume harinas, número de días que realiza actividad física, mediciones en el análisis de hemoglobina: leucositos, eritrositos, hemoglobina, ematrocitos, MCV, MCH, MCHC, RDW, plaquetas, MPV, neutrófilos, linfocitos, monocitos, eosinófilos, basófilos, neutrofilos AVS, linfocitos AVS, monofositos AVS, eosinófilos AVS, tipo de sangre, glucosa en ayuno. Mediciones en el análisis de orina donde se obtuvieron urea, creatinina sérica, ácido úrico, densidad, PH y células epiteliales.

Figura 4.7 Especificando el modelo en R[®]

```
"model {
  for( i in 1 : N ) {
    Condicion[i] ~ dbern(mu[i])
    mu[i] <- 1 / (1 + exp(-(b0 + b1*Edad[i] + b2*Peso[i] + b3*Estatura[i] + b4*PPM[i] + b5*IVS[i]
    + b6*CS[i] + b7*Edad CA + b8*Dias sangrado[i] + b9*Embarazos[i] + b10*Cesareas[i] +
    b11*Partos[i] + b12*Abortos[i] + b13*duerme[i] + b14*Consume maiz + b15*Carnes ro-
    jas[i] + b16*Harinas[i] + b17*act.Fisica[i])))
  }
  b0 ~ dnorm( 0 , 1.0E-12 )
  b1 ~ dnorm( 0 , 1.0E-12 )
  b2 ~ dnorm( 0 , 1.0E-12 )
  b3 ~ dnorm( 0 , 1.0E-12 )
  b4 ~ dnorm( 0 , 1.0E-12 )
  b5 ~ dnorm( 0 , 1.0E-12 )
  b6 ~ dnorm( 0 , 1.0E-12 )
  b7 ~ dnorm( 0 , 1.0E-12 )
  b8 ~ dnorm( 0 , 1.0E-12 )
  b9 ~ dnorm( 0 , 1.0E-12 )
  b10 ~ dnorm( 0 , 1.0E-12 )
  b11 ~ dnorm( 0 , 1.0E-12 )
  b12 ~ dnorm( 0 , 1.0E-12 )
  b13 ~ dnorm( 0 , 1.0E-12 )
  b14 ~ dnorm( 0 , 1.0E-12 )
  b15 ~ dnorm( 0 , 1.0E-12 )
  b16 ~ dnorm( 0 , 1.0E-12 )
  b17 ~ dnorm( 0 , 1.0E-12 )
  }", file="logistic.jag")
```

Comenzamos especificando un modelo. El proceso es similar al de un modelo lineal. Establecimos un ciclo sobre cada observación, pero aquí especificamos la probabilidad como un resultado distribuido por Bernoulli (en lugar de distribuido normalmente) con el parámetro μ definido explícitamente como un modelo logístico. Después de especificar la probabilidad, especificamos distribuciones previas para los coeficientes β . Aquí solo aplicamos distribuciones normales vagamente informativas para permitir que los datos hablen por sí mismos; esto se observa en la figura 4.7. Después cargamos el conjunto de datos y definimos la estructura de datos para la ejecución de JAGS®. Se debe tener en cuenta que estamos incluyendo el número de filas para la variable N en nuestra declaración de modelo.

Primero es necesario cargar los datos en el programa R® y verificar si se cargaron para después declarar listas y poder iniciar a trabajar con ellos. En la figura 4.8 se muestran las instrucciones de cómo realizarlo en el programa R®:

Figura 4.8 Comandos para cargar datos en R®

```
#Cargamos los datos
Datos=read.delim("clipboard",header=T,stringsAsFactors=F)
#Los verificamos
attach(Datos)
```

| X0 | Edad | Peso | Estatura | PPM | IVS | CS | Edadca | sangrado | Embarazos | Cesareas |
|----|------|------|----------|-----|-----|----|--------|----------|-----------|----------|
| 0 | 22 | 46 | 1.74 | 13 | 16 | 5 | 24 | 7 | 1 | 0 |
| 0 | 57 | 82 | 1.56 | 12 | 28 | 1 | 55 | 0 | 0 | 0 |
| 0 | 33 | 74 | 1.64 | 13 | 23 | 2 | 33 | 5 | 1 | 1 |
| 0 | 47 | 90 | 1.64 | 11 | 16 | 10 | 46 | 4 | 3 | 0 |
| 0 | 21 | 50 | 1.56 | 12 | 20 | 1 | 26 | 4 | 0 | 0 |

| Partos | Abortos | duerme | maiz | Carnes | Harinas | Fisica |
|--------|---------|--------|------|--------|---------|--------|
| 0 | 1 | 7 | 6 | 6 | 6 | 3 |
| 0 | 0 | 8 | 5 | 3 | 3 | 5 |
| 0 | 0 | 8 | 4 | 1 | 0 | 5 |
| 3 | 0 | 7 | 4 | 1 | 6 | 7 |
| 0 | 0 | 8 | 2 | 5 | 5 | 2 |

```
# Se declaran la lista (son tres maneras)
dat1<-list("Condicion", "Edad", "Peso", "n")
```

```
dat1<-list("Condicion", "Edad", "Peso", "n")
dat1<-list(Condicion=Datos$Condicion, Edad=Datos$Edad,
Peso=Datos$Peso, Estatura=Datos$Estatura, N=nrow(Da-
tos))
Condicion <- Datos$Condicion
Edad <- Datos$Edad
Peso <- Datos$Peso
Estatura<-Datos$Estatura
```

Al establecer nuestros valores iniciales, adoptamos un enfoque diferente. La maquinaria de MCMC, los valores iniciales pueden tener un efecto sobre la eficiencia con la que el algoritmo cubre el espacio muestral. No existen reglas estrictas y rápidas para elegir valores iniciales más allá de la recomendación de que sean realistas (para que el muestreador MCMC funcione de manera eficiente) y sobredispersos (para que podamos evaluar mejor la convergencia a una distribución posterior estable). Un enfoque es establecer la elección de los valores iniciales en una simple ejecución de EMV (estimación de máxima verosimilitud) del modelo. Esto tiene el beneficio adicional de permitir comparar sus resultados bayesianos con los de un enfoque frecuentista. Aquí, extraemos los coeficientes de un modelo logístico MLG (modelo lineal generalizado) (para obtener valores realistas) y sumamos/restamos un número razonable (para dispersar los valores iniciales). La estructura del objeto de datos que contiene los valores iniciales es una lista de listas. Todo lo anterior se observa en la figura 4.9 aplicado en R®:

Figura 4.9 Estableciendo valores iniciales y realización de MLG

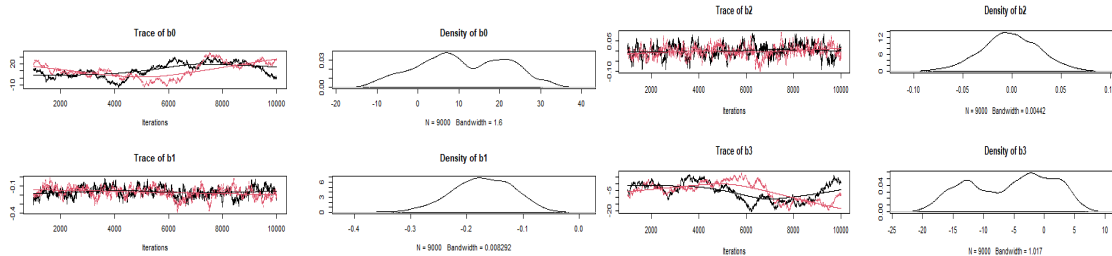
```
n=length(Condicion)
data_mas<-list("Condicion", "Edad", "Peso", "Estatura", "PPM", "IVS", "CS",
"EdadCA", "sangrado", "Embarazos", "Cesáreas", "Partos",
"Abortos", "duerme", "maiz", "rojas", "Harinas", "física", "n")
param<-c("b0", "b1", "b2", "b3", "b4", "b5", "b6", "b7", "b8", "b9", "b10",
"b11", "b12", "b13", "b14", "b15", "b16", "b17")
estInits<-with(Datos, glm(Condicion~Edad+Peso+Estatura+PPM+IVS+CS+EdadCA+sangrado+Embarazos+Ce-
sareas+Partos+Abortos+duerme+maiz+rojas+Harinas+física, family=binomial(logit)))
estInits
Call: glm(formula = X0 ~ Edad + Peso + Estatura + PPM + IVS + CS +
Edadca + sangrado + Embarazos + Cesareas + Partos + Abortos +
duerme + maiz + Carnes + Harinas + Física, family = binomial(logit))
Coefficients:
(Intercept) Edad Peso Estatura PPM IVS CS
7.5166948 -0.1238300 -0.0005422 -3.9870618 -0.0066538 0.2494619 0.0798821
Edadca sangrado Embarazos Cesareas Partos Abortos duerme
0.0177384 0.0760389 -5.0163397 5.0028905 5.5628617 5.5000412 -0.3686576
maiz Carnes Harinas Física
-0.0617891 -0.1384231 0.1100823 -0.0409853
Degrees of Freedom: 71 Total (i.e. Null); 54 Residual
Null Deviance: 99.81
Residual Deviance: 76.36 AIC: 112.4
```

Las estimaciones puntuales son muy cercanas a las de la ejecución MLG. Se esperaba esto. Cuando tiene una gran cantidad de datos y antecedentes muy vagos o solo mínimamente informativos, la probabilidad predomina y obtiene las estimaciones EMV. Por supuesto, todavía se tienen los beneficios de poder realizar evaluaciones de probabilidad directa de sus resultados, que no están disponibles en un modelo estricto frecuentista MLE, y puede hacer cosas como actualizar sus resultados, agregar observaciones, calcular probabilidades para funciones complejas, etc.

Se realizan las gráficas de las MCMC para determinar si existe convergencia y se puede observar la instrucción y las gráficas de salida en R[®] en la figura 4.10.

Figura 4.10 Gráficas MCMC de los primeros tres parámetros

`plot(as.mcmc(logmod))`

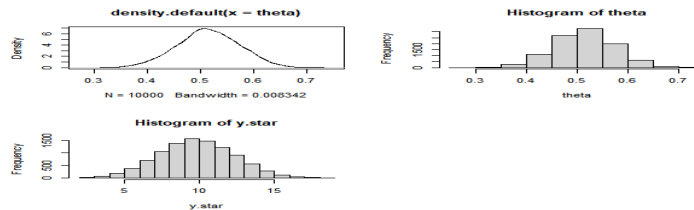


La validación predictiva consiste en verificar qué tan bien los datos generados a partir de su modelo coinciden con las observaciones reales. Esto se puede hacer a través de predicciones externas, reteniendo algunos datos (digamos 10 %) al ajustar su modelo y comparar esas observaciones con las generadas al conectar valores de predictores en su modelo. O puede usar todos sus datos para ajustar el modelo y comparar los resultados observados con los predichos por el modelo.

Es posible crear un conjunto de datos predictivos posteriores para cada individuo u observación en su conjunto de datos mediante el uso de simulaciones que incluyen la varianza adicional que caracteriza los análisis bayesianos. Los residuos se calculan restando los valores simulados de los valores observados. Las parcelas residuales se evalúan como suelen ser. Sin embargo, en un enfoque bayesiano, podemos repetir este proceso muchas veces simulando conjuntos de datos predictivos y calculando residuos basados en ellos. Entonces, por ejemplo, si le preocupa la heterocedasticidad, puede calcular la correlación entre el valor absoluto de los residuos y los valores ajustados muchas veces, y trazar un histograma de las muchas correlaciones que calcula. En la figura 4.11 se muestran las instrucciones y el histograma en R[®]:

Figura 4.11 Histograma para determinar heterocedasticidad

```
theta<-rbeta(10000,(sum(Condicion)+1),((sum(n)-sum(Condicion))-1))  
plot(density(theta))  
hist(theta)  
y.star<-rbinom(10000,20, theta)  
hist(y.star)
```



5. Análisis de resultados

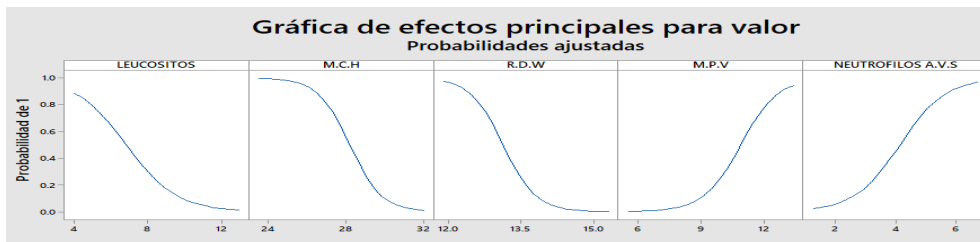
Una vez realizado el análisis, los factores que fueron significativos y que explican la evolución de VPH en el modelo de regresión logística bayesiana que se desarrolló fueron los siguientes: entre 21-29 años de edad, entre 1.59-1.64 metros de estatura, tipo de sangre O⁺, primer periodo menstrual haya sido entre los 12-13 años de edad, su vida sexual inició entre los 19-22 años, si ha tenido entre 1-4 compañeros sexuales y estos tienen entre 23-31 años de edad o entre 41-49 años de edad. Además, que sus periodos menstruales duren entre 3-5 días, ha estado embarazada 1 o 2 veces, si no ha tenido cesáreas, partos ni abortos. Si duerme entre 7 y 8 horas diarias, si consume maíz de 6 a 7 días por semana, si consume carnes rojas de 2 a 3 días por semana o consume harinas de 6 a 7 días por semana.

Con las variables del historial clínico referentes a los análisis clínicos tanto de biometría hemática y general de orina, los factores significativos fueron: leucocitos, MCH (concentración de hemoglobina corpuscular media, mide el contenido de hemoglobina en los eritrocitos), RDW (la amplitud de la distribución eritrocitaria [ADE o RDW, en sus siglas en inglés] o ancho de distribución de glóbulos rojos, MPV (mide el tamaño promedio de las plaquetas) y neutrófilos AVS (son un tipo de glóbulo blanco y ayudan al cuerpo a combatir infecciones) como se muestra la figura 5.1 y 5.2.

Figura 5.1 Análisis de regresión logística bayesiana

| Tabla de desviaciones | | | | | |
|-----------------------|----|--------------|--------------|--------------|---------|
| Fuente | GL | Desv. ajust. | Media ajust. | Chi-cuadrada | Valor p |
| Regresión | 19 | 60.0694 | 3.1615 | 60.07 | 0.000 |
| LEUCOSITOS | 1 | 3.9279 | 3.9279 | 3.93 | 0.047 |
| ERITROSITOS | 1 | 0.0330 | 0.0330 | 0.03 | 0.856 |
| HEMOGLOBINA | 1 | 1.2448 | 1.2448 | 1.24 | 0.265 |
| EMATROCITOS | 1 | 1.0964 | 1.0964 | 1.10 | 0.295 |
| M.C.V | 1 | 2.2201 | 2.2201 | 2.22 | 0.136 |
| M.C.H | 1 | 8.8528 | 8.8528 | 8.85 | 0.003 |
| M.C.H.C | 1 | 0.6337 | 0.6337 | 0.63 | 0.426 |
| R.D.W | 1 | 13.2068 | 13.2068 | 13.21 | 0.000 |
| PLAQUETAS | 1 | 0.0104 | 0.0104 | 0.01 | 0.919 |
| M.P.V | 1 | 6.4639 | 6.4639 | 6.46 | 0.011 |
| NEUTROFILOS | 1 | 1.5680 | 1.5680 | 1.57 | 0.210 |
| LINFOSITOS | 1 | 2.3109 | 2.3109 | 2.31 | 0.128 |
| MONOCITOS | 1 | 0.3398 | 0.3398 | 0.34 | 0.560 |
| EOSINOFILOS | 1 | 1.4307 | 1.4307 | 1.43 | 0.232 |
| BASOFILOS | 1 | 1.5453 | 1.5453 | 1.55 | 0.214 |
| NEUTROFILOS A.V.S | 1 | 9.8708 | 9.8708 | 9.87 | 0.002 |
| LINFOSITOS A.V.S | 1 | 0.1992 | 0.1992 | 0.20 | 0.655 |
| MONOFOSITOS A.V.S | 1 | 0.2027 | 0.2027 | 0.20 | 0.653 |
| EOSINOFILOS A.V.S | 1 | 0.5085 | 0.5085 | 0.51 | 0.476 |
| Error | 52 | 39.7438 | 0.7643 | | |
| Total | 71 | 99.8132 | | | |

Figura 5.2 Gráfica de probabilidades ajustadas



Con base en la hipótesis establecida, los factores significativos en las pacientes con VPH que se plantearon fueron los siguientes: edad, peso, estatura, primer periodo menstrual, inicio de vida sexual, compañeros sexuales, compañeros circuncidados, duración de periodo menstrual, embarazos, estatus socioeconómico y vida clínica; además, si con el análisis bayesiano era posible desarrollar un sistema que modelara la extensión del virus. Así, fue posible establecer que los factores significativos son la edad, el inicio de vida sexual, los días que consume maíz, los leucocitos, MCH, RDW, MPV y los neutrofilos AVS. Además, se logró desarrollar el modelo para determinar el diagnóstico del VPH mediante inferencia bayesiana, que se muestra en la ecuación 5.1 (los factores significativos se muestran en las figuras 5.3 y 5.4).

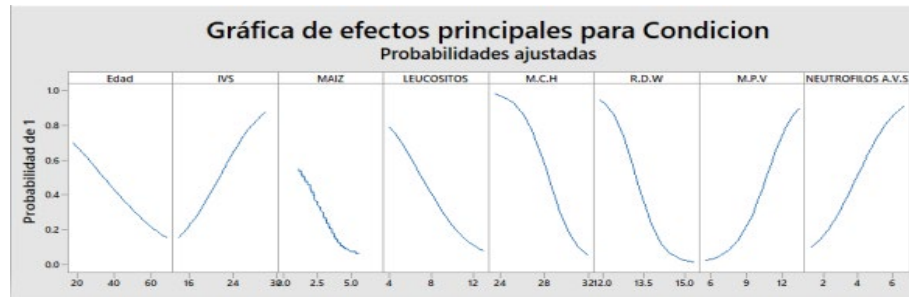
El modelo resultante fue:

$$Y' = 42.1 - 0.0482Edad + 0.229IVS - 0.318MAIZ - 0.420LEUCOSITOS - 0.801M.C.H - 2.217R.D.W + 0.782M.P.V + 0.844NEUTROFILOS A.V.S \quad (5.1)$$

Figura 5.3 Resultados de regresión logística bayesiana para los factores principales

| Fuentes | GL | Desv. Ajust. | Media Ajust. | Chi-cuadrada | Valor p |
|-------------------|----|--------------|--------------|--------------|---------|
| Regresión | 7 | 50.638 | 7.2340 | 50.64 | 0.000 |
| Edad | 1 | 1.900 | 1.8996 | 1.90 | 0.168 |
| IVS | 1 | 5.126 | 5.1262 | 5.13 | 0.024 |
| MAIZ | 1 | 4.6147 | 4.61 | 4.61 | 0.032 |
| Leucocitos | 1 | 3.362 | 3.3616 | 3.36 | 0.067 |
| M.C.H | 1 | 13.312 | 12.3116 | 12.31 | 0.000 |
| R.D.W | 1 | 16.150 | 16.1497 | 16.15 | 0.000 |
| M.P.V | 1 | 7.897 | 7.8968 | 7.90 | 0.005 |
| NEUTROFILOS A.V.S | 1 | 7.656 | 7.6568 | 7.66 | 0.006 |
| ERROR | 64 | 49.175 | 0.7684 | | |
| Total | 71 | 99.813 | | | |

Figura 5.4 Gráfica de efectos principales de los factores del modelo final



Conclusiones

El objetivo fundamental de esta investigación era generar un modelo para determinar la confiabilidad del diagnóstico del VPH mediante inferencia bayesiana, clave para demostrar que con inferencia bayesiana es posible predecir con base en evidencia existente. Quizás lo más convincente es que ofrece una manera atractiva e intuitiva pero metodológicamente rigurosa de actualizar nuestro conocimiento. Así pues, la aportación principal de este trabajo consiste dar a conocer la inferencia bayesiana y el modelo que logramos generar, el cual se muestra en el capítulo cinco, después de un proceso que inició tocando puertas para tener acceso a la información para posteriormente realizar campañas de concientización, estudio de colposcopia y revisando los estudios de biometría y general de orina. De este modo, se lograron identificar los factores predominantes del VPH, donde los 44 factores analizados se dividían en factores físicos y de comportamiento: 17 y 27 de análisis clínicos, solo 8 resultaron significativos en el modelo final.

Cabe destacar que uno de esos factores significativos fue si la paciente consumía tortillas de maíz entre 6 y 7 días de la semana, y en un análisis realizado por Meléndez *et al.* (2020), que lleva por título *Análisis de aceites esenciales como bioconservadores*,

a través de la teoría de confiabilidad se encontró que el conservador químico utilizado en las tortillas de maíz evidencia que provoca cáncer. Por ello, resulta interesante ahondar en dicho factor.

Referencias

- Ashby, D., Hutton, J. L. and McGee, M. A. (1993). Simple Bayesian analyses for case-control studies in cancer epidemiology. *Journal of the Royal Statistical Society*, 42(4), 385-397.
- Berumen-Campos, J. (2003). Nuevos virus del papiloma humano descubiertos en México: Su asociación a la alta incidencia del cáncer de cérvix. *Gaceta Médica de México*, 139(s4), 3-10.
- Centro Nacional de Equidad de Género y Salud Reproductiva (2007). Seguimiento a los programas y líneas de acción para la atención de la salud de las mujeres del Centro Nacional de Equidad de Género y Salud Reproductiva (CNEGSR) de la Secretaria de Salud.
<https://www.gob.mx/salud/cnegsr>
- Correa, J. C. y Barrera, C. J. (2018). *Introducción a la estadística bayesiana: notas de clase*. Instituto Tecnológico Metropolitano.
- Flores Sánchez, A. (2006). *Aplicación de herramientas de ingeniería (Mahalanobis-Taguchi) para identificar los factores de riesgo del VPH*. Instituto Tecnológico de Ciudad Juárez.
- Flores-Miramontes, M. G., Torres-Reyes, L. A., Aguilar-Lemarroy, A., Vallejo-Ruiz, V., Piña-Sánchez, P., Cortés-Gutiérrez, E. y Jave-Suarez, L. F. (2015). Prevalencia de genotipos de VPH en México y en el mundo detectadas mediante Linear Array. *Revista Médica del Instituto Mexicano del Seguro Social*, 53(2), S122-S130.
- Instituto Nacional de Estadística Geografía e Informática (INEGI) (2007). *Anuario Estadístico de los Estados Unidos Mexicanos*.
- Meléndez, Y (2020). *Análisis de aceites esenciales como bioconservadores, a través de la teoría de confiabilidad* (tesis doctoral). Instituto Tecnológico de Ciudad Juárez, México.
- Muñoz, N., Bosch, F., Sanjosé, S., Viladiu, P., Tormo, J., Moreo, P., Ascunce, N., González, L., Tafur, L., Gili, M., Izarzugaza, I., Guerrero, E., Aristizábal, N., Santamaría, M., Alonso de Ruiz, P. y Shah, H. (1993). El virus del papiloma humano en la etiología del cáncer cervicouterino. *Boletín de la Oficina Sanitaria Panamericana (OPS)*, 115(4).
- Rivera, Z., Aguilera, J. y Larraín, A. (2002). Epidemiología del virus papiloma humano (HPV). *Revista Chilena de Obstetricia y Ginecología*, 67, 501-506.
- Serman, F. (2002). Cáncer cervicouterino. Epidemiología, historia natural y rol del virus del papiloma humano. Perspectivas en prevención y tratamiento. *Revista Chilena Infectología*, 26(2), 174-190.
- Stanczuk, G. A., Kay, P., Allan, B., Chirara, M., Tswana, S. A., Bergstrom, S. and Williamson, A. L. (2003). Detection of human papillomavirus in urine and swabs from patients with invasive cervical cancer. *Journal of Medical Virology*, 71(1), 110-114.
- Zamudio, A. A., Zepeda, Z. J., Rodríguez, B. B. y Tenorio, M. R. (2001). Evaluación del papanicolaou y la colposcopia en el diagnóstico de la infección por el virus del papiloma humano. *Rev, Fac MedUNAM*, 44(1), 5-7.