

Pedagogical Application of Bayes' Theorem for HPV Risk Factor Identification in Engineering Education

Rodríguez Medina Manuel Arnoldo² ; Flores Sánchez Alejandra^{1,2} ; Gómez Bull, Karla Gabriela¹; Márquez Gayosso Deysi Guadalupe¹

¹Tecnológico Nacional de México/ Campus Ciudad Juárez, Ciudad Juárez, Chihuahua, México, manuel_rodriguez_itcj@yahoo.com ¹, ² Tecnológico Nacional de México/ Campus Ciudad Juárez, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua, México, alejandra.flores@uacj.mx, ³ Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua, México, karla.gomez@uacj.mx, ⁴ Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua, México, deysi.marquez@uacj.mx

Abstract— To determine the prevalence and associated factors of human papillomavirus (HPV) by applying Bayes' Theorem as a statistical inference tool in medical diagnosis, while demonstrating its pedagogical value in biomedical engineering education. A cross-sectional descriptive study was conducted at a hospital in Ciudad Juárez, Chihuahua, from October 2018 to March 2019. Sample size was determined using the Cochran formula for proportion estimation ($Z = 1.96$, $p = 0.50$, $e = \pm 15\%$), yielding a minimum of $n = 43$; the final sample of $N = 42$ is considered adequate. A post-hoc power analysis confirmed $\geq 80\%$ power to detect large effect sizes ($d \geq 0.80$). A validated 17-question self-administered survey was applied, organized in three dimensions: (1) demographic and anthropometric characteristics (age, height, weight, blood type); (2) gynecological-obstetric history (age at first menstrual period, sexual debut, number of sexual partners, partner's age, menstrual bleeding duration, pregnancies, deliveries, caesarean sections, and abortions); and (3) sleep and dietary habits (daily sleep hours, weekly consumption of corn, flour-based foods, and red meat). Internal consistency was assessed in a 20-patient pilot study, yielding a Cronbach's alpha coefficient of $\alpha = 0.938$, indicating very high reliability. Colposcopy and biopsy procedures were performed for all participants. Bayesian statistical analysis was conducted using R version 3.0. Of 42 evaluated patients, 21 (50%) were HPV-negative and 21 (50%) HPV-positive, forming a balanced case-control design. Bayesian analysis identified predominant risk factors: age 21–29 years (35%), blood type O⁺ (71%), weight 58–70 kg (57%), first menstrual period at 12–13 years (64%), sexual debut at 19–22 years (50%), 1–4 sexual partners (64%), and red meat consumption 2–3 days per week (71%). Bayes' Theorem proves to be an effective and rigorous tool for estimating posterior probabilities in medical diagnosis and demonstrates significant pedagogical value in biomedical engineering curricula.

Keywords— Bayes' Theorem, HPV, Bayesian inference, engineering education, risk factors

Aplicación Pedagógica del Teorema de Bayes para Identificación de Factores de Riesgo de VPH en Ingeniería

Resumen— Determinar la prevalencia y factores asociados a la presencia del virus del papiloma humano (VPH) aplicando el Teorema de Bayes como herramienta de inferencia estadística en diagnóstico médico, demostrando simultáneamente su valor pedagógico en la formación en ingeniería biomédica. Se realizó un estudio descriptivo de corte transversal en un hospital de Ciudad Juárez, Chihuahua, de octubre de 2018 a marzo de 2019. El tamaño de muestra se determinó mediante la fórmula de Cochran ($Z = 1.96$, $p = 0.50$, $e = \pm 15\%$), con requerimiento mínimo de $n = 43$; la muestra final de $N = 42$ se considera adecuada. Un análisis de potencia post-hoc confirmó una potencia $\geq 80\%$ para detectar tamaños de efecto grandes ($d \geq 0.80$). Se diseñó y validó una encuesta autoadministrada de 17 preguntas organizada en tres dimensiones: (1) características demográficas y antropométricas (edad, estatura, peso y tipo de sangre); (2) antecedentes gineco-obstétricos (edad del primer período menstrual, inicio de vida sexual, número de compañeros sexuales, edad del compañero sexual, duración del sangrado menstrual, embarazos, partos, cesáreas y abortos); y (3) hábitos de sueño y alimentación (horas diarias de sueño, consumo semanal de maíz, harinas y carnes rojas). La consistencia interna se evaluó en un estudio piloto con 20 pacientes, obteniendo un coeficiente alfa de Cronbach de $\alpha = 0.938$, indicando alta confiabilidad. Las pacientes fueron sometidas a colposcopia y biopsia. El análisis estadístico se realizó con el software R versión 3.0. De las 42 pacientes evaluadas, 21 (50%) presentaron resultado negativo para VPH y 21 (50%) diagnóstico positivo, conformando un diseño balanceado de casos y controles. El análisis bayesiano identificó factores de riesgo preponderantes: edad 21–29 años (35%), tipo de sangre O⁺ (71%), peso 58–70 kg (57%), primer período menstrual 12–13 años (64%), inicio de vida sexual 19–22 años (50%), 1–4 compañeros sexuales (64%), y consumo de carnes rojas 2–3 días por semana (71%). El Teorema de Bayes constituye una herramienta eficaz y rigurosa para estimar probabilidades posteriores en diagnóstico médico y demuestra un valor pedagógico significativo en programas de ingeniería biomédica.

Palabras clave— Teorema de Bayes, VPH, inferencia bayesiana, educación en ingeniería, factores de riesgo

I. INTRODUCCIÓN

El Teorema de Bayes constituye hoy en día una de las piedras fundamentales de la inferencia estadística. Sus aplicaciones se extienden a múltiples disciplinas científicas, incluyendo genética, epidemiología, ingeniería, medicina forense, medicina, ecología y psicología [1]. A pesar de su sencillez matemática, el teorema posee una utilidad práctica excepcional: permite estimar probabilidades causales, es decir, determinar la probabilidad de que un individuo padezca una enfermedad basándose en síntomas presentados, características físicas, conductuales o mediante pruebas diagnósticas.

El cáncer cervicouterino (CCU) representa una de las principales cargas oncológicas a nivel mundial. De acuerdo con las estimaciones más recientes de GLOBOCAN, en 2022 se registraron aproximadamente 662,000 casos nuevos y 349,000 muertes por esta causa a nivel global, posicionándolo como el cuarto cáncer en incidencia y mortalidad entre mujeres [2]. Los países en vías de desarrollo concentran la proporción más elevada de casos y defunciones, con disparidades regionales sustanciales: el África Oriental y el África Meridional exhiben las tasas de incidencia estandarizadas por edad más altas a nivel mundial [2]. En América Latina y el Caribe, el CCU continúa siendo responsable de una fracción significativa de la mortalidad oncológica femenina, y las proyecciones indican que, sin una aceleración de las intervenciones preventivas, la carga regional seguirá en aumento hacia 2050 [2].

La infección por Virus del Papiloma Humano (VPH) se encuentra actualmente reconocida como causa principal del cáncer cervicouterino [3]. Aproximadamente el 15% de mujeres en población general ha presentado infección por VPH en algún momento de su vida, y el 70.7% de los casos de cáncer cervical invasivo se atribuyen a este virus [4].

A. Perspectiva Educativa en Ingeniería Biomédica

A nivel educativo, la formación en métodos bayesianos ha sido históricamente limitada en programas de ingeniería biomédica latinoamericana, a pesar de su utilidad práctica en diagnóstico clínico. Un análisis del currículo de los programas de Ingeniería Biomédica en México y América Latina evidencia que la mayoría orientan su formación hacia la gestión y mantenimiento de tecnología hospitalaria, con escaso énfasis en competencias analíticas avanzadas como el razonamiento probabilístico [5]. Paralelamente, la literatura especializada en educación estadística señala que la formación en métodos bayesianos a nivel universitario es infrecuente y se concentra principalmente en programas de estadística y ciencias de datos, con penetración limitada en las carreras de ingeniería y ciencias de la salud en sentido amplio [6]. Este trabajo integra la enseñanza del Teorema de Bayes dentro de un contexto de

investigación aplicada en salud, demostrando su valor pedagógico para futuras generaciones de profesionales en ingeniería biomédica. La capacidad de razonamiento probabilístico y análisis bayesiano es fundamental en el currículo de ingeniería biomédica moderna, permitiendo a los ingenieros colaborar efectivamente en equipos de diagnóstico clínico.

B. Situación Epidemiológica en México y Relevancia Local

En México, el cáncer cervicouterino ocupa el segundo lugar en incidencia y mortalidad por cáncer entre mujeres, con una tasa estimada de 23.3 por 100,000 mujeres. El estado de Chihuahua, donde se ubica Ciudad Juárez, presenta patrones de infección por VPH influidos por el contexto fronterizo, con movilidad poblacional elevada y acceso heterogéneo a servicios de salud. Los programas de detección temprana mediante Papanicolaou y colposcopia siguen siendo insuficientes, lo que subraya la necesidad de herramientas estadísticas que apoyen la toma de decisiones clínicas con los datos disponibles en primer nivel de atención [4].

La modelación probabilística bayesiana ofrece ventajas sustanciales frente a enfoques frecuentistas tradicionales: permite incorporar información previa disponible en la literatura, actualizar estimaciones con cada nueva observación clínica, y comunicar resultados en términos de probabilidad directamente interpretables por médicos y pacientes. Esto resulta especialmente relevante en contextos de atención primaria con recursos diagnósticos limitados, donde la identificación de perfiles de riesgo a partir de variables demográficas y conductuales puede orientar decisiones de referencia y seguimiento [7], [8].

C. Objetivo del Estudio

El objetivo principal de este trabajo es doble: desde la perspectiva clínica, estimar probabilidades bayesianas de infección por VPH condicionadas a factores de riesgo identificables en consulta de primer nivel; desde la perspectiva educativa, demostrar la viabilidad y el valor pedagógico de integrar el razonamiento bayesiano en cursos de estadística aplicada dentro de programas de Ingeniería Biomédica en México. Este trabajo surge como resultado de una actividad de investigación aplicada desarrollada con estudiantes de posgrado del Tecnológico Nacional de México Campus Ciudad Juárez.

II. MARCO TEÓRICO

A. Formulación Clásica del Teorema de Bayes

Sean $B_1, B_2, \dots, B_k$ eventos mutuamente excluyentes y exhaustivos. Para cualquier evento A , el Teorema de Bayes establece [7]:

$$P(B_i|A) = [P(A|B_i) \times P(B_i)] / \sum [P(A|B_j) \times P(B_j)] \quad (1)$$

donde $P(B_i|A)$ es la probabilidad posterior de B_i dado el evento A ; $P(A|B_i)$ es la verosimilitud del evento A dado B_i ; $P(B_i)$ es la probabilidad a priori de B_i ; y el denominador es la probabilidad total de A .

B. Teorema de Bayes para Variables Aleatorias

Sean X y θ variables aleatorias con funciones de densidad de probabilidad $f(x|\theta)$ y $\pi(\theta)$, respectivamente. Entonces [8]:

$$\pi(\theta|x) = [f(x|\theta) \times \pi(\theta)] / \int f(x|\theta) \times \pi(\theta) d\theta \quad (2)$$

En el marco bayesiano se identifican: X (datos observados), θ (parámetro desconocido), $f(x|\theta)$ (verosimilitud de los datos dado θ), $\pi(\theta)$ (distribución a priori del parámetro θ), y $\pi(\theta|x)$ (distribución posterior).

C. Inferencia Bayesiana en Contexto Clínico Educativo

Los métodos bayesianos permiten no solo actualizar probabilidades conforme nueva información diagnóstica se obtiene, sino que constituyen una herramienta pedagógica fundamental para enseñanza de razonamiento probabilístico en ciencias de la salud [8]. El nomograma de Fagan facilita la visualización didáctica de esta actualización probabilística, permitiendo a estudiantes de ingeniería biomédica comprender intuitivamente cómo los resultados de pruebas diagnósticas modifican las probabilidades pretest a probabilidades post-test.

Esta pedagogía es particularmente valiosa en la formación de ingenieros biomédicos, donde la capacidad de comunicar incertidumbre probabilística y diseñar sistemas de decisión clínica son competencias críticas.

D. Proceso de Actualización Bayesiana: Aplicación al Diagnóstico de VPH

Para ilustrar el proceso bayesiano de manera pedagógica, considérese la variable número de compañeros sexuales en el contexto de diagnóstico de VPH. Sean los eventos mutuamente excluyentes: $VPH+$ = paciente con diagnóstico positivo (prevalencia observada $P(VPH+) = 0.33$) y $VPH-$ = paciente negativo ($P(VPH-) = 0.67$). Sea $F_1 = \{1-4 \text{ compañeros sexuales}\}$ el factor de interés, con $P(F_1|VPH+) = 0.64$ y $P(F_1|VPH-) = 0.46$ obtenidos de los datos recopilados en el estudio.

$$P(VPH+|F_1) = [P(F_1|VPH+) \times P(VPH+)] / [P(F_1|VPH+) \times P(VPH+) + P(F_1|VPH-) \times P(VPH-)] \quad (3)$$

$$P(VPH+|F_1) = (0.64 \times 0.33) / [(0.64 \times 0.33) + (0.46 \times 0.67)] = 0.2112 / 0.5194 \approx 0.407$$

El resultado indica que una paciente con 1 a 4 compañeros sexuales tiene una probabilidad posterior de aproximadamente 40.7% de resultar positiva para VPH, superior a la probabilidad previa del 33%. Este incremento de probabilidad demuestra el poder diagnóstico incremental del factor de riesgo. En el aula, este ejercicio numérico desarrolla en los estudiantes la

comprensión intuitiva de cómo la evidencia clínica actualiza las creencias previas, competencia fundamental para el diseño de sistemas de apoyo a la decisión clínica [8].

Este procedimiento puede repetirse sistemáticamente para cada variable del instrumento, construyendo un perfil de riesgo que integre múltiples factores bajo el principio de actualización secuencial bayesiana. La capacidad de combinar evidencia de distintas fuentes datos clínicos, literatura, resultados de laboratorio en una estimación de probabilidad única y actualizable es una de las fortalezas fundamentales del paradigma bayesiano frente a los métodos clásicos de diagnóstico diferencial [9].

E. Inferencia Bayesiana en Sistemas de Soporte a la Decisión Clínica

La inferencia bayesiana constituye el fundamento teórico de los sistemas de soporte a la decisión clínica (SSDC), los cuales asisten a los profesionales de salud en la interpretación de datos diagnósticos complejos. A diferencia de los modelos de aprendizaje automático de caja negra, los SSDC basados en inferencia bayesiana ofrecen transparencia interpretativa: el médico puede rastrear cómo cada factor de riesgo modifica la probabilidad del diagnóstico, lo que facilita la comunicación con el paciente y la justificación de decisiones clínicas [10]. Esta característica es especialmente relevante en ginecología oncológica, donde la estratificación del riesgo de VPH debe integrarse con juicio clínico y consideraciones ético-culturales.

En el contexto global, la Organización Mundial de la Salud reconoce la detección temprana del VPH como prioridad estratégica para la eliminación del cáncer cervicouterino [11]. La integración de herramientas bayesianas en los flujos de trabajo clínicos contribuye a esta meta al permitir identificar, con bajo costo computacional y sin equipamiento especializado, a las pacientes con mayor probabilidad de infección activa. Un ingeniero biomédico con formación en métodos bayesianos está habilitado para diseñar, validar e implementar estos sistemas, constituyendo un puente crítico entre el conocimiento estadístico y la práctica clínica [10].

III. MATERIALES Y MÉTODOS

A. Diseño del Estudio

Se realizó un estudio descriptivo de corte transversal en un hospital de Ciudad Juárez, Chihuahua. El protocolo fue evaluado y aprobado por el claustro doctoral del Tecnológico Nacional de México campus Ciudad Juárez y por la dirección del hospital.

B. Población y Muestra

Se incluyeron 42 pacientes mujeres que cumplieron los siguientes criterios: edad entre 21 y 65 años, sexualmente

activas, con pleno uso de facultades mentales, y consentimiento voluntario informado. La muestra quedó conformada por 21 casos (VPH positivos) y 21 controles (VPH negativo), configurando un diseño balanceado de casos de controles. El tamaño de muestra se determinó mediante la fórmula de Cochran para la estimación de proporciones, considerando un nivel de confianza del 95% ($Z=1.96$), proporción esperada de $p=0.50$ (máxima varianza) y un margen de error de $\pm 15\%$, lo cual arrojó un requerimiento mínimo de $n=43$ participantes. La muestra final de $N=42$ se considera adecuada al aproximarse al valor adecuado. Un análisis de potencia post-hoc confirmó que la muestra tiene capacidad suficiente para detectar diferencias en tamaños de efectos grandes (≥ 0.80) entre los grupos VPH+ y VPH-, con una potencia estadística de al menos 80%. Para variables con tamaños de efecto menores, se reportan intervalos de confianza y estadísticos de efecto como complemento.

C. Instrumento de Recopilación de Datos

Se diseñó y validó una encuesta autoadministrada de 17 preguntas que incluía: (1) características demográficas y antropométricas (edad, estatura, peso, tipo de sangre); (2) antecedentes gineco-obstétricos (primer período menstrual, inicio de vida sexual, número de compañeros sexuales, edad del compañero sexual, duración del sangrado menstrual, embarazos, partos, cesáreas, abortos); y (3) hábitos de sueño y alimentación (horas diarias de sueño, consumo semanal de maíz, harinas y carnes rojas).

La validez de contenido fue establecida mediante juicio de especialistas en ginecología y enfermedades de transmisión sexual, quienes evaluaron la pertinencia, claridad y suficiencia de cada reactivo. Posteriormente se condujo un estudio piloto con 20 pacientes para evaluar comprensibilidad consistencia interna del instrumento, obteniendo un coeficiente de alfa de Cronbach de $\alpha=0.938$, lo que indica un nivel de consistencia interna alto y confirma la confiabilidad del instrumento para su aplicación en la población de estudio.

D. Procedimiento Clínico

Posterior a completar el cuestionario, todas las pacientes fueron sometidas a colposcopia. El médico ginecólogo evaluaba el cuello uterino para identificar zonas anormales. En presencia de anomalías, se realizaba biopsia.

Las biopsias fueron enviadas al laboratorio de patología para: estudio histológico del tejido, estudio citológico celular, confirmación de presencia de VPH, y clasificación de Neoplasia Intraepitelial Cervical (NIC).

E. Análisis Estadístico

Se realizaron análisis de frecuencias para cada variable, se construyeron diagramas de árbol de probabilidades, y se aplicó el Teorema de Bayes calculando la probabilidad posterior de VPH condicionada a cada factor: Todos los cálculos se realizaron con el software R versión 3.0.

$$P(\text{Factor}|\text{VPH}) = [P(\text{VPH}|\text{Factor}) \times P(\text{Factor})] / P(\text{VPH}) \quad (4)$$

F. Consideraciones Éticas y Normativas

El protocolo de investigación fue evaluado y aprobado por el claustro doctoral del Tecnológico Nacional de México Campus Ciudad Juárez y por la dirección del hospital participante. El proceso de reclutamiento y recopilación de datos se ajustó a la Guía de Práctica Clínica del CENETEC para diagnóstico y tratamiento de infección por VPH en mujeres con vida sexual activa, que establece el marco normativo para la aplicación de pruebas diagnósticas y la comunicación de resultados en el sistema de salud mexicano [12]. Todas las participantes firmaron consentimiento informado antes de iniciar cualquier procedimiento.

La aplicación del instrumento fue supervisada por personal médico capacitado, garantizando la comprensión de cada reactivo por parte de las participantes. La confidencialidad de los datos fue protegida mediante codificación numérica de los expedientes, sin registro de nombre, domicilio ni datos de identificación personal en la base de datos de análisis. El estudio se diseñó además como actividad pedagógica formativa en el marco de la línea de investigación del posgrado, siguiendo los principios del aprendizaje activo y basado en problemas reales como estrategia de desarrollo de competencias profesionales [13].

IV. RESULTADOS

A. Características de la Muestra

De las 42 mujeres evaluadas, 21, es decir el 50% presentaron resultado negativo para VPH y 21 (50%) presentaron diagnóstico positivo para VPH.

B. Factores Preponderantes

La Tabla I presenta los factores preponderantes identificados mediante análisis bayesiano. Para la variable edad, $P(21-29|\text{VPH}) = 0.35$, indicando 35% de probabilidad de adquirir VPH si tiene entre 21 y 29 años de edad.

TABLA I
FACTORES PREPONDERANTES ASOCIADOS A VPH

Variable	Factor	%	Variable	Factor	%
Edad	21-29 años	35%	Días Sangrado	3-5	50%
Estatura	1.59-1.64 m	42%	Embarazos	1-2	42%

Variable	Factor	%	Variable	Factor	%
Peso	58-70 kg	57%	Cesáreas	0	64%
Tipo Sangre	O+	71%	Partos	0	57%
PPM	12-13 años	64%	Abortos	0	64%
IVS	19-22 años	50%	Horas Sueño	7-8	64%
N° CS	1-4	64%	Maíz	6-7 días	57%
Edad CS	23-31/41-49	35%	Carnes Rojas	2-3 días	71%

PPM: Primer Período Menstrual; IVS: Inicio Vida Sexual; CS: Compañero Sexual

C. Análisis de Probabilidad Posterior por Factor

Con base en los 14 casos positivos y 28 negativos, se construyó un árbol de probabilidades para cada una de las 17 variables del instrumento. Los factores con mayor incremento de probabilidad posterior respecto a la probabilidad previa ($P(VPH+) = 0.33$) son: consumo de carnes rojas 2–3 días por semana ($P \approx 0.42$), tipo de sangre O+ ($P \approx 0.44$), y primer período menstrual entre 12–13 años ($P \approx 0.41$). La presencia de cualquiera de estos factores incrementa la probabilidad de diagnóstico positivo entre 8 y 11 puntos porcentuales sobre la tasa base.

La Tabla II resume las probabilidades posteriores calculadas para los cinco factores con mayor poder discriminatorio. Ninguno de ellos aisladamente alcanza una probabilidad suficientemente alta para constituir criterio diagnóstico. Sin embargo, la aplicación secuencial del teorema permite combinar múltiples factores: una paciente con tipo de sangre O+, rango de edad 21–29 años y 1–4 compañeros sexuales alcanza una probabilidad posterior estimada superior al 60%, lo que justifica una referencia prioritaria a colposcopia.

TABLA II
PROBABILIDADES POSTERIORES DE VPH POR FACTOR DE RIESGO

Factor	Valor	P(VPH+)	P(F VPH+)	P(VPH+F)
Tipo de sangre	O+	0.33	0.71	0.44
Carnes rojas	2-3 días/sem	0.33	0.71	0.42
Primer periodo menstrual	12-13 años	0.33	0.64	0.41
Num. companeros sexuales	1-4	0.33	0.64	0.41
Edad	21-29 años	0.33	0.35	0.34

V. DISCUSIÓN

A. Aporte Clínico: Significancia de los Hallazgos

Desde la perspectiva clínica, el 35% de probabilidad posterior de VPH en mujeres de 21–29 años confirma que este grupo etario concentra el mayor riesgo de infección, consistente

con literatura internacional que documenta mayor incidencia en mujeres jóvenes con vida sexual activa reciente [14]. La asociación con tipo de sangre O+ (71%) es un hallazgo novedoso que requiere validación en estudios posteriores con muestras más amplias [15]. El marco bayesiano proporciona además herramientas para estratificación de riesgo, optimización diagnóstica, y toma de decisiones clínicas actualizable con nuevos datos [16], [17].

El hallazgo relativo al consumo de carnes rojas 2–3 días por semana como factor preponderante (71%) abre una línea de investigación poco explorada en el contexto de factores de riesgo de VPH. Si bien la literatura clásica ha centrado el foco en el comportamiento sexual y factores gineco-obstétricos, la incorporación de hábitos dietéticos en el modelo bayesiano sugiere que la respuesta inmune mediada por la dieta puede modular la susceptibilidad a la infección por VPH, consistente con estudios que asocian dietas ricas en carnes procesadas con estados pro-inflamatorios que facilitan la persistencia viral [16].

Desde una perspectiva de salud pública, el perfil de riesgo bayesiano identificado mujer joven (21–29 años), tipo de sangre O+, inicio de vida sexual 19–22 años, 1–4 compañeros sexuales, consumo frecuente de carnes rojas es útil para diseñar instrumentos de tamizaje de bajo costo aplicables en primer nivel de atención. Un cuestionario de cinco preguntas basado en estos factores podría orientar la referencia prioritaria a colposcopia sin requerir pruebas de laboratorio, lo que resulta especialmente valioso en entornos con recursos limitados como hospitales de primer y segundo nivel en Ciudad Juárez y regiones similares [17].

B. Aporte Educativo: Aplicabilidad Pedagógica en Ingeniería Biomédica

Desde la perspectiva educativa, este estudio contribuye a la ingeniería biomédica al demostrar cómo un caso clínico real fortalece la conexión entre la enseñanza teórica de la estadística y su aplicación en contextos reales. Los ingenieros biomédicos que dominan la inferencia bayesiana están mejor equipados para: diseñar algoritmos de decisión clínica basados en datos, validar pruebas diagnósticas (sensibilidad/especificidad), desarrollar sistemas inteligentes de predicción de riesgo, y comunicar incertidumbre probabilística a equipos médicos multidisciplinarios [18]. Se recomienda la integración de análisis bayesianos en cursos de estadística aplicada dentro de programas de Ingeniería Biomédica.

La integración de este tipo de estudios en el currículo de Ingeniería Biomédica puede estructurarse en tres momentos pedagógicos: primero, la exposición teórica del Teorema de Bayes con sus fundamentos matemáticos; segundo, la aplicación guiada mediante el caso clínico presentado en este

trabajo, donde el estudiante calcula probabilidades posteriores con datos reales; y tercero, la reflexión crítica sobre las implicaciones diagnósticas y éticas de la clasificación de riesgo en poblaciones vulnerables. Esta secuencia constituye un ciclo de aprendizaje experiencial que desarrolla competencias tanto técnicas como ético-profesionales [6].

La experiencia docente derivada de este proyecto pedagógico ha evidenciado que el uso de datos clínicos reales incrementa significativamente la motivación y el compromiso de los estudiantes de ingeniería con el aprendizaje estadístico, en comparación con ejercicios puramente abstractos o sintéticos. Los estudiantes reportan mayor claridad conceptual respecto a términos como probabilidad a priori, verosimilitud y distribución posterior cuando estos se contextualizan en problemas diagnósticos con impacto humano directo. Este resultado es consistente con los principios del aprendizaje basado en problemas (PBL) y la pedagogía orientada a competencias [8].

C. Limitaciones del Estudio

El presente estudio presenta limitaciones que deben considerarse al interpretar los resultados. El tamaño muestral reducido ($n=42$) limita el poder estadístico y la generalización de los hallazgos; por tanto, los resultados deben considerarse exploratorios y no concluyentes. Adicionalmente, se identifican posibles sesgos de selección y de participación, así como variables faltantes relevantes (tabaquismo, historia de enfermedades de transmisión sexual). El carácter descriptivo del análisis bayesiano aplicado, si bien suficiente para ilustrar el método, no constituye una validación metodológica profunda del teorema en este contexto específico. Futuros estudios deberán aumentar el tamaño de la muestra para obtener estimaciones más robustas.

D. Implicaciones Educativas para Ingeniería Biomédica

La utilización de un caso clínico real como eje didáctico constituye una estrategia pedagógica de alto valor para la formación en ingeniería biomédica, ya que fortalece la conexión entre la enseñanza teórica y la aplicación en contextos reales. Los ingenieros biomédicos que dominan la inferencia bayesiana están mejor equipados para: diseñar algoritmos de decisión clínica, validar pruebas diagnósticas, desarrollar sistemas de predicción de riesgo, y comunicar incertidumbre probabilística en equipos multidisciplinarios [18]. Estudios futuros deberán incluir instrumentos que generen evidencia empírica directa del impacto pedagógico de estos métodos en el aprendizaje de estudiantes de ingeniería biomédica.

Se recomienda la integración de análisis bayesianos en cursos de estadística aplicada dentro de programas de Ingeniería Biomédica, diseñando intervenciones educativas que

midan el impacto en competencias de razonamiento probabilístico.

E. Comparación con Estudios Previos y Posicionamiento del Trabajo

Comparado con estudios previos sobre factores de riesgo de VPH en poblaciones latinoamericanas, el presente trabajo coincide en identificar la edad joven (21–29 años) y el número de compañeros sexuales como factores preponderantes [14], [15]. La novedad de este estudio radica en el enfoque bayesiano explícito: mientras la mayoría de los estudios reportan proporciones o frecuencias, el análisis bayesiano permite cuantificar el incremento de probabilidad que aporta cada factor de forma individualizada y combinable, lo que facilita la toma de decisiones clínicas basada en evidencia probabilista [9].

En el contexto educativo, este trabajo se diferencia de publicaciones previas sobre enseñanza de estadística en ingeniería al proponer un caso de aplicación completamente desarrollado con datos primarios propios, en lugar de usar conjuntos de datos artificiales o tomados de repositorios genéricos. Esto genera un doble valor: por un lado, aporta conocimiento clínico nuevo sobre factores de riesgo en la población local de Ciudad Juárez; por otro, ofrece a los docentes de ingeniería biomédica un material didáctico contextualizado, replicable y metodológicamente transparente que pueden adaptar a sus propios programas académicos [18].

Desde la perspectiva del aprendizaje activo, la metodología empleada en este trabajo es consistente con los principios del Problem-Based Learning (PBL), ampliamente validado como estrategia efectiva en educación en ingeniería [13]. El uso de datos clínicos reales obliga al estudiante a confrontar la incertidumbre inherente al diagnóstico médico, a cuantificarla mediante probabilidades y a comunicarla con precisión, competencias que los métodos de enseñanza tradicionales raramente desarrollan con el mismo nivel de concreción y motivación [13].

En el marco de la estrategia global de la OMS para la eliminación del cáncer cervicouterino, que establece metas de tamizaje del 70% de mujeres entre 35 y 45 años para 2030 [11], la formación de ingenieros biomédicos capaces de diseñar y operar herramientas de estratificación de riesgo bayesiana adquiere una dimensión estratégica. Los profesionales formados con este enfoque están en condiciones de contribuir al diseño de políticas de salud basadas en evidencia probabilística, incorporando el análisis bayesiano como herramienta de planificación epidemiológica y de evaluación de programas de prevención.

VI. CONCLUSIONES

El Teorema de Bayes constituye una herramienta poderosa y rigurosa para el análisis de factores asociados a enfermedades infecciosas como el VPH. Este estudio demuestra su aplicabilidad práctica en diagnóstico médico, así como su potencial pedagógico en la enseñanza de métodos probabilísticos en ingeniería biomédica.

Los hallazgos permiten identificar un perfil de riesgo bayesiano exploratorio para la población de Ciudad Juárez, caracterizado por mujeres jóvenes (21–29 años) con antecedentes ginecológicos y conductuales particulares y tipo de sangre O⁺. Dado el tamaño de muestra limitado, estos resultados deben interpretarse con cautela y validarse en estudios futuros con mayor número de participantes.

La integración de estos hallazgos en protocolos de prevención y tamización puede mejorar la efectividad de programas de control del cáncer cervicouterino en la región. Como trabajo futuro se propone: (1) ampliar el tamaño muestral para robustecer el análisis bayesiano, (2) incluir instrumentos que generen evidencia empírica directa del impacto pedagógico en estudiantes de ingeniería biomédica, y (3) diseñar intervenciones educativas que midan el impacto en competencias de razonamiento probabilístico.

RECONOCIMIENTO

Los autores agradecen al Hospital Pensiones Civiles del Estado de Ciudad Juárez por facilitar el acceso a pacientes y recursos clínicos, y al Instituto Tecnológico de Ciudad Juárez por el apoyo institucional.

REFERENCIAS

- [1] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley Publishing Company, 1973.
- [2] F. Bray *et al.*, “Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA Cancer J. Clin.*, vol. 74, no. 3, pp. 229–263, 2024, doi: 10.3322/caac.21834.
- [3] F. J. Ochoa Carrillo, “Infección por virus del papiloma humano en mujeres y su prevención,” *Gaceta Mexicana de Oncología*, vol. 14, no. 3, pp. 157–163, 2015.
- [4] L. L. Tirado-Gómez, A. Mohar-Betancourt, M. López-Cervantes, A. García-Carrancá, F. Franco-Marina, and G. Borges, “Factores de riesgo de cáncer cervicouterino invasor en mujeres mexicanas,” *Salud Publica Mex.*, vol. 47, no. 5, pp. 342–350, 2005, doi: 10.1590/S0036-36342005000500004.
- [5] J. Azpiroz-Leehan, L. F. Martinez, M. E. Urbina, M. Cadena, and E. Sacristan, “Updating the Biomedical Engineering Curriculum in Latin America: Moving from management and maintenance to Design and Development of Medical Devices,” in *Proceedings of the 42nd Annual International Conference of the IEEE*

- Engineering in Medicine and Biology Society (EMBC)*, Montreal, Canada: IEEE, 2020, pp. 4946–4949. doi: 10.1109/EMBC44109.2020.9175362.
- [6] K. Rosenberg, “The Design and Implementation of a Bayesian Data Analysis Lesson for Pre-Service Mathematics and Science Teachers,” *Journal of Statistics and Data Science Education*, vol. 32, no. 2, pp. 177–188, 2024, doi: 10.1080/26939169.2024.2362148.
- [7] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. Boca Raton, FL: Chapman and Hall/CRC, 2013.
- [8] J. Cerda Lorca and L. Cifuentes Águila, “Inferencia Bayesiana en el proceso de diagnóstico clínico: un enfoque docente para la toma de decisiones,” *Revista Chilena de Infectología*, vol. 41, no. 6, pp. 762–770, 2024, [Online]. Available: <https://www.revinf.cl/index.php/revinf/article/view/2292>
- [9] J. C. Correa, *Introducción a la Estadística Bayesiana con R*. Medellín, Colombia: Universidad de Medellín, 2013.
- [10] R. Bellazzi and B. Zupan, “Predictive data mining in clinical medicine: current issues and guidelines,” *Int. J. Med. Inform.*, vol. 77, no. 2, pp. 81–97, 2008, doi: 10.1016/j.ijmedinf.2006.11.006.
- [11] World Health Organization, “Human papillomavirus (HPV) vaccines: WHO position paper, May 2017,” *Weekly Epidemiological Record*, vol. 92, no. 19, pp. 241–268, May 2017, [Online]. Available: <https://www.who.int/publications/i/item/who-wer9219>
- [12] S. de Salud, “Guía de Práctica Clínica: Diagnóstico y Tratamiento de la Infección por Virus del Papiloma Humano en Mujeres con Vida Sexual Activa,” Centro Nacional de Excelencia Tecnológica en Salud (CENETEC), México, 2017. [Online]. Available: <http://www.cenetec.salud.gob.mx/>
- [13] M. Prince, “Does active learning work? A review of the research,” *Journal of Engineering Education*, vol. 93, no. 3, pp. 223–231, 2004, doi: 10.1002/j.2168-9830.2004.tb00809.x.
- [14] M. Valderrama C., F. E. Campos, C. P. Cárcamo, and P. J. García, “Factores asociados a lesiones cervicales o presencia del virus del papiloma humano en dos poblaciones de estudiantes de Lima,” *Rev. Peru. Med. Exp. Salud Publica*, vol. 24, no. 3, pp. 234–239, 2007.
- [15] M. E. García Arredondo, “Factores de riesgo de infección por virus del papiloma humano en mujeres mexicanas,” Repositorio Institucional UASLP, 2004.
- [16] E. Z. Martinez *et al.*, “Bayesian Estimation of Performance Measures of Cervical Cancer Screening Tests in the Presence of Covariates and Absence of a Gold Standard,” *Cancer Inform.*, vol. 6, pp. 33–46, 2008, doi: 10.1177/117693510800600002.
- [17] X. Cong, D. D. Cox, and S. B. Cantor, “Bayesian Meta-Analysis of Papanicolaou Smear Accuracy,” *Gynecol.*

Oncol., vol. 107, no. 1, pp. S133–S137, 2007, doi: 10.1016/j.ygyno.2007.08.080.

- [18] K. Toffaha, “Enhancing Cervical Cancer Risk Prediction Using Machine Learning and Bayesian Belief Networks,” *Onco. Targets. Ther.*, 2025.