

Análisis de la relación entre hipertensión arterial y algunos hábitos de la población

Brandon Stif Campos Abarca¹, Sebastián Pacheco Picado¹, Kevin Andrey Montenegro Barboza¹

brandon.camposabarca@ucr.ac.cr, sebastian.pachecopicado@ucr.ac.cr,
kevin.montenegrobarboza@ucr.ac.cr

RESUMEN

La medición de la hipertensión arterial puede ser bastante compleja si no se cuenta con los aparatos tecnológicos necesarios, sin embargo, dentro de las poblaciones humanas se debe conocer si la hipertensión arterial es un padecimiento frecuente que aumenta la mortalidad. Por lo que se necesita crear una ecuación que estime la presión arterial promedio. Con la selección de variables asociadas a un aumento en la presión arterial se generó un modelo inicial, teniendo como variables edad, peso, altura, cantidad de fumado, consumo de comida procesada, consumo de alcohol, la raza, familiares cercanos a la unidad estadística (UE) que padezcan hipertensión arterial y si la UE consume algún tratamiento para que altere la presión arterial. A partir de aquí, se observaron supuestos importantes, tales como, normalidad, homocedasticidad, linealidad e independencia. Dentro de los supuestos no se pudo remediar la heteroscedasticidad, por lo que se generó un modelo que asume heteroscedasticidad, pero utilizando el método de mínimos cuadrados ponderados. Posteriormente se realizó la selección de variables del modelo, mediante el método de la prueba Fisher con selección de variables hacia atrás. Por último, se analizó el porcentaje de la explicación del modelo seleccionado, el cual contuvo únicamente dos variables que lograron explicar un 31% de la variabilidad ponderada de la Presión arterial las cuales fueron peso y comida.

PALABRAS CLAVE: Presión arterial, VIF, mínimos cuadrados ponderados, modelo de regresión lineal.

INTRODUCCIÓN

Uno de los objetivos de las investigaciones médicas es postergar la muerte de la población. Sin embargo, en los últimos 40 años las enfermedades crónicas no transmisibles han aumentado considerablemente (Espinosa, 2012, p.51). Con la inserción de la industrialización en la población, la expectativa de vida cambió abruptamente; el sedentarismo se normalizó con rapidez y con esto los malos hábitos alimenticios, entre otros (Calvo, 2010, p.3). Dentro de estas prácticas, uno de los efectos adversos en el ser humano es la hipertensión arterial. Casado (2002) define la presión arterial como “la sangre impulsada por el corazón fluye por el torrente circulatorio, o mejor, por el sistema arterial, sometida a una presión denominada presión arterial (PA), o tensión arterial” (p.121), por medio de esto la enfermedad de hipertensión arterial es definida como “[...] una elevación de la presión arterial, en la que se han excluido las causas secundarias (feocromocitoma, aldosteronismo, insuficiencia renal y otras), así como las formas mendelianas (monogénicas)” (Carretero, Oparil,

¹ Estudiante de la Universidad de Costa Rica

Messeri, Williams & Ritz, citados por Calvo, 2010, p.60), además Cohn y Col (citados por Calvo, 2010) explican la hipertensión arterial como “Se trata de un estado de funcionamiento anormal y de cambios estructurales de las arterias asociado con disfunción endotelial, vasoconstricción o remodelación del músculo liso vascular; aumento en la impedancia del ventrículo izquierdo y propensión a la aterosclerosis. [...]” (p.60). Entonces se propone un análisis de la presión arterial mediante un modelo de regresión lineal múltiple, donde la variable respuesta (presión arterial) se explicará por medio de algunos hábitos y características de la población que alteren o perjudiquen a esta, y así poder identificar aquellas poblaciones que están padeciendo hipertensión arterial.

Con ello, existe una necesidad de crear una ecuación de regresión lineal que, con características de recolección inmediatas, de un acercamiento al promedio de la presión arterial y con ello determinar si las poblaciones pueden padecer de hipertensión arterial. Para crear el modelo de regresión lineal múltiple se valoraron varias variables independientes que estuviesen relacionadas a la hipertensión arterial y que estas mismas alteren la presión arterial.

El consumo de comida procesada puede tener relación con la hipertensión arterial; como lo explica Huertas el consumo de grasas y grasas saturadas encontradas en alimentos de origen animal eleva los niveles de colesterol en la sangre, así como el consumo de sodio (2011, p.209). Igualmente, la raza puede relacionarse,

Estudios longitudinales han demostrado que la raza negra es la de mayor incidencia, pero actualmente por los cambios en el ritmo de vida y la no modificación de los factores de riesgo está aumentando la incidencia en las demás etnias (González, 1999 & Balaguer, 1998, citados por Huertas, 2011, p.209).

Algunos artículos mencionan la relación de las variables edad de las personas y sexo como factor de la hipertensión arterial, de acuerdo con Huerta “El ser varón es un factor de riesgo para cardiopatía isquémica e hipertensión arterial” (2001, p.209) y agrega a su vez que entre los 35 a 40 años el hombre tiene una mortalidad cuatro veces mayor que en la mujer (Huertas, 2001, p.209). Dando una buena referencia a estas variables de estudio, se mencionan también investigaciones donde se habla que antes de los 40 años el riesgo es mayor sin embargo luego de los 60 años tiende a tener un comportamiento similar, influenciado por los hábitos de las personas y su desarrollo “La relación de hipertensión/sexo puede ser modificada por la edad” (Guevara, Ranero, Álvarez, Álvarez, y Tápanes, 2022, p.5). Dentro de los hábitos mencionados se destaca el consumo de tabaco, el cual incrementa y acelera el daño vascular, influye en las paredes de las arterias por lo tanto en la circulación de estas y agrava la resistencia a la insulina, entre otros. Huerta, (2001, p.209).

Además, la hipertensión arterial puede tener relación con el alcohol pues “El alcohol puede producir una elevación aguda de la presión arterial mediada por activación simpática central cuando se consume en forma repetida y puede provocar una elevación persistente”, lo cual puede llevar a hipertensión crónica si no se regula (Huerta, 2001, p.209). Por otro lado, el peso también es una variable para considerar, la obesidad está claramente asociada con un mayor riesgo de hipertensión, ya que un aumento en el índice de masa corporal (IMC) incrementa significativamente las posibilidades de desarrollar hipertensión, especialmente en poblaciones jóvenes (Álvarez, Ramírez-Campillo, Martínez-Salazar, Vallejos-Rojas, Jaramillo-Gallardo, Salas Bravo, Cano-Montoya & Celis-Morales, 2016). Por último, la actividad física regular disminuye la rigidez de las arterias e

incrementa la sensibilidad a la insulina, ayudando a reducir los niveles de presión arterial, especialmente en personas con sobrepeso (Huerta, 2001, p.209).

Estas características podrían generar un buen acercamiento al promedio real de la hipertensión arterial con base a estas variables, puesto que en muchas ocasiones la medición de la hipertensión arterial puede no ser de fácil acceso para todas las personas; se enfoca el objetivo general: desarrollar una ecuación lineal que prediga el promedio de la hipertensión arterial usando variables con un alto grado de explicación. Con ello, en forma específica, un análisis del comportamiento de la hipertensión arterial promedio en relación con factores que influyen en este.

Dado estos objetivos; se generó la hipótesis de investigación, donde se cuestiona si alguna de las variables seleccionadas logra predecir el promedio de la hipertensión arterial de forma confiable.

METODOLOGÍA

El trabajo se realizó en la Universidad de Costa Rica, ubicada en San Pedro de Montes de Oca, en la provincia de San José, en el año 2024, con una población aproximada de 54458 personas (estudiantes y trabajadores) (Suplemento OCI, Universidad de Costa Rica, 2024).

La Universidad de Costa Rica está subdividida en facultades (Letras, Económicas, Medicina, Farmacia, Artes, Estudios Generales, Aulas, Matemáticas, Ciencias Sociales, Computacional, Geografía, Ciencias, Odontología, Biología, Agronomía, Química y Música) por constante, la recolección de datos se realizó por el método de conglomerados. Se realizó una captura de 77 unidades estadísticas, donde la unidad estadística es un estudiante o trabajador activo de la Universidad de Costa Rica en el año 2024, con una edad en años cumplidos mayor a 18 y menor a 70 años. Primeramente, se eligieron diez facultades de forma aleatoria, luego de ello, se escogieron aproximadamente 10 unidades estadísticas de forma aleatoria por cada facultad que se encontraron en el momento de recolección dentro de la facultad seleccionada. Ya elegida la unidad estadística se procedió a recolectar la información de esta por medio de dos fases; primero, se le cuestionó a la Unidad Estadística sobre si tomó cafeína o realizó alguna actividad física una hora antes de la captura, de ser positiva la respuesta, se seleccionó otra UE (Unidad estadística). De ser negativa la respuesta, se le dio un reposo de 10 minutos antes de la toma de datos, luego se procedió a la toma de la muestra de la presión arterial mediante el uso de Esfigmomanómetro automático (dispositivo que permite medir la presión sanguínea, el método de uso es por medio electrónico, se colocó en el brazo izquierdo de la UE, unos dos o tres dedos por encima del plexo, debía quedar a la altura del corazón, los resultados obtenidos fueron la presión sistólica “arriba” de la presión diastólica, ambas en mmHg). Después se utilizó una balanza calibrada para la toma del peso, donde la UE no usó zapatos ni objetos pesados en su cuerpo, y para terminar la primera fase se midió la altura con una cinta métrica que se colocó en un sector con una inclinación de 90° aproximadamente, poniendo una regla en la coronilla de la cabeza haciendo un ángulo de 90° entre la coronilla y la cinta métrica para obtener la estatura. Para la segunda fase, se le pidió a la UE que contestara un cuestionario, donde se capturaron por medio de preguntas las demás variables (Cuadro 1, anexo). Al finalizar con esto, se obtuvo la captura completa de los datos.

Para la creación del modelo se tomó en cuenta como variable respuesta la presión arterial (mmHg), como en el proceso de la captura de los datos se registraron los valores de la presión arterial sistólica y diastólica por aparte, se debió utilizar la presión arterial media (PAM), esto dado

que el experto Fallas, explicó que para unir ambas variables se podía utilizar la técnica descrita. Además, si la PAM daba mayor a 100 se consideraba que esa persona padecía hipertensión arterial. Ahora bien, las variables independientes correspondían al peso (kilogramos), la altura (centímetros), si la UE tiene algún familiar cercano en su línea sanguínea que padezca de hipertensión arterial, la cantidad de fumado diario (sea marihuana, vapor o cigarrillo), la cantidad de consumo de comida procesada diario, el sexo de la UE, la cantidad de días que realiza actividad física por semana, el tipo de raza (mestizo o blanco), la cantidad de consumo de bebidas alcohólicas semanales y si la UE consume algún tratamiento que contrarrestare efectos hipertensivos.

Dado las variables; se definió el modelo lineal:

$$\mu_{X|Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \alpha_1 X_1 X_9 + \alpha_2 X_2 X_9$$

Donde;

X_1 = Edad

X_2 = Peso

X_3 = Altura

X_4 = Fuma

X_5 = Comida

X_6 = Alcohol

X_7 = Raza (Categórica; blanca=0, mestiza=1)

X_8 = Parentesco (Categórica; Sí=1, No=0)

X_9 = Sexo (Categórica; Hombre= 0, Mujer=1)

X_{10} = Tratamiento (Categórica; Sí= 1, No= 0)

Por otro lado, en la Figura 2 (anexo) se muestra el diagrama de flujo, donde se observan los pasos que se llevaron a cabo para realizar la investigación, esta manera visual de presentar el proceso ayudó a organizar las ideas y revisar los pasos previos antes de llegar a una conclusión. En primer lugar, se definió el tema junto a sus objetivos y las variables propuestas, se realizó la recolección de los datos (77 Unidades Estadísticas) y a su vez se definió el modelo inicial. Luego de esto, se realizó la prueba de interacción para la selección de variables comparando el modelo inicial con el modelo con interacción para conocer si la interacción es significativa, a raíz de este proceso ya el modelo elegido, se analizaron diagnósticos de valores extremos y los valores de influencia sobre las estimaciones.

A partir de aquí, se realizaron distintas pruebas de supuestos (normalidad, homocedasticidad, linealidad, multicolinealidad). En algunos casos como la homocedasticidad se debió utilizar medidas remediales como mínimos cuadrados ponderados. Luego de todo este proceso, se realizó una selección por pasos hacia adelante y nuevamente se realizó una validación de supuestos (normalidad, linealidad, multicolinealidad). Una vez que se validó el modelo, se procedió a darle la escritura del modelo y a la ecuación de este mismo. A este modelo final, se le analizó el porcentaje de explicación mediante el coeficiente de determinación R^2 .

El análisis se realizó por medio de gráficos y diferentes pruebas con el lenguaje de R con la versión 4.2.2 (R Core Team, 2022), utilizando los siguientes paquetes: car (Fox, Weisberg, 2019) para construir gráficos como el de cuantil-cuantil, así como distintos análisis de dispersión, lattice (Sarkar,

Deepayan, 2008) para graficar y analizar interacciones, así como las relaciones de las variables predictoras con la respuesta.

Además, como experto y colaborador de la investigación se contó con Daniel Francisco Fallas Álvarez, estudiante avanzado de la carrera de Enfermería, de la Universidad de Costa Rica, y como experta y colaboradora de la investigación se contó con Rosa Gabriela Granados Vargas, enfermera de la Caja Costarricense del Seguro Social.

RESULTADOS

Primeramente, se realizó un análisis exploratorio de las variables predictoras numéricas pretendiendo conocer en primera instancia estadísticos descriptivos.

Cuadro 1

Análisis Descriptivo de Variables Predictoras Numéricas

Variables	Estadísticos descriptivos			
	Mínimo	Media	Máximo	Desviación Estándar
Edad	18	20	54	6.83
Peso	46.95	69.87	113.75	14.39
Altura	1.69	1.49	1.88	0.093
Alcohol	0	0.28	2	0.58
Comida	0	3.122	10	2.36
Fuma	0	0.28	5	0.96

El Cuadro 1 muestra que la variable peso se dispersó entre 46.95 y 113.75, demostrando la diversidad de pesos que se capturaron dentro de la muestra seleccionada. En relación con la edad, se dispersó entre los 18 y 54 años, siendo una de las variables más complicadas al momento de recolección de los datos, puesto que edades más altas era complejo de capturar por el lugar de selección de la muestra. Luego de esto, otra variable importante de observar es la de Fuma, pues la media de fumado de las personas de la Universidad de Costa Rica es casi cero.

Se observó la relación existente entre las variables predictoras y la variable respuesta. En la Figura 4 (anexo), la relación de la respuesta con edad mantiene una forma lineal, sin embargo, es casi constante (se observa un aumento poco pronunciado), al igual que con la variable fuma y alcohol. En cambio, la relación de la presión con el peso se muestra que conforme aumenta el peso, la presión aumenta en promedio. Al igual que con la variable altura, la relación muestra que conforme la altura aumenta, la presión en promedio va a aumentar. Y con respecto a la comida, conforme aumenta la cantidad de consumo de comida procesada, la presión en promedio va a aumentar.

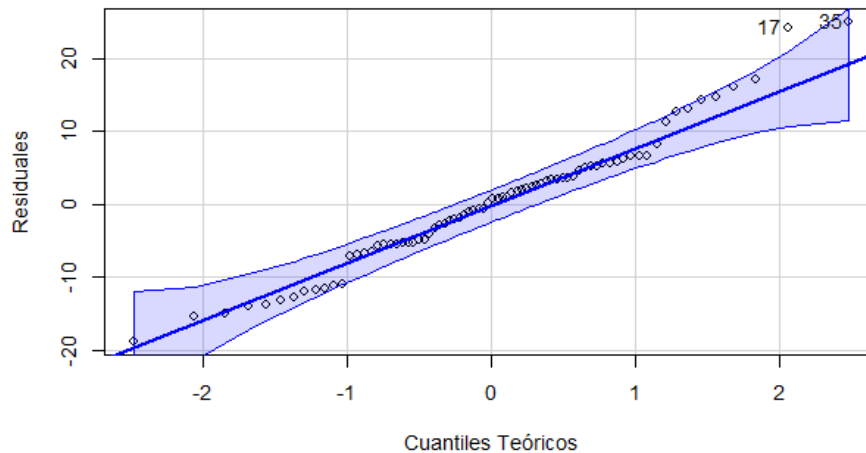
En seguida, se hizo el estudio de los valores extremos, con base en los residuales estudentizados, y no se observó presencia de estos. Por otra parte, se estudió a los valores de influencia sobre los valores ajustados mediante la distancia de Cook pero no hubo existencia de valores de influencia.

Analizando las variables categóricas, se realizó la visualización en la Figura 5 (anexo) de las interacciones que se podrían generar con algunas variables numéricas. Visualmente se podría determinar que hay existencia de interacción entre edad y sexo, así como entre comida y

tratamiento, pero se realizó la comparación, del modelo propuesto con un modelo que no tuviera interacción, mediante la prueba Fisher. Con una probabilidad asociada a 0.7585 se procedió a no rechazar la hipótesis nula de que los coeficientes de las variables con interacción eran iguales a cero. A partir de esto, se continuó con los demás análisis a partir del modelo sin interacción.

Figura 1.

Distribución de los residuales del modelo contra los cuantiles teóricos.



En cuanto al supuesto de la normalidad se realizó un gráfico de los residuales contra los cuantiles teóricos, para observar si la distribución condicional de la variable respuesta contra las variables predictoras es normal. Como se observa en la Figura 1, los residuales se encuentran dentro de las bandas de confianza, por ello se puede suponer que el modelo cumple con el supuesto de normalidad. Además, para realizar la confirmación del supuesto, se realizó la prueba de Shapiro Wilks donde se obtuvo una probabilidad asociada de 0.1659, con esto no se rechaza la hipótesis nula que los residuales pueden seguir una distribución normal, por constante, se asume normalidad en el modelo.

En relación con la homocedasticidad primero se realizó un gráfico de los residuales de modelo contra los valores ajustados donde en la Figura 6 (anexo) sí se puede observar una forma de patrón. Esto se confirma, al realizar la prueba formal de Breusch Pagan se obtuvo una probabilidad baja ($p < 0.001$), rechazando la hipótesis nula de que la varianza condicional de la variable respuesta es igual para cada valor de los predictores, con esto se asumió que no hay homocedasticidad. Por lo tanto, se procedió a realizar una transformación a la variable respuesta para ayudar a corregir el problema de la no heterocedasticidad, se analizó la transformación de Box-cox. Como se observa en la Figura 7 (anexo), dentro de los valores que maximizan la verosimilitud de la respuesta se encuentra el cero, se procedió a transformar a la variable respuesta con un logaritmo natural y generar un modelo con la variable respuesta transformada. Al analizar la prueba formal de Breusch Pagan se obtuvo una probabilidad de 0.126 no rechazando la hipótesis nula antes mencionada.

A partir del nuevo modelo que contiene el logaritmo, se observa el supuesto de linealidad. Como se observa en la Figura 8 (anexo), se podría pensar que, en el caso de edad, mantienen una curva bastante pronunciada, sin embargo, al aplicarle alguna medida remedial, esto no cambia, y sigue siendo no significativa. Por lo que se decidió dejarla de esa forma. Luego, en el caso de las variables peso, altura y comida, se observan bastante lineales. Con ello, se procedió a realizar el análisis de la multicolinealidad, donde se procedió a calcular el Factor de Inflación de la Variancia

(VIF), pues este va a medir la cantidad en la que la variancia de un coeficiente se va a ver inflado dado que los predictores puedan estar correlacionados. Pero se demostró que el VIF para ninguna de las variancias de los coeficientes de las variables predictoras sobrepasa a 10, asumiendo que el modelo no posee multicolinealidad.

Con los supuestos ya remediados y analizados, se procedió a seleccionar las variables. En este caso, se utilizó la selección por pasos hacia atrás con la prueba. Al momento de quitar aquellas variables que no eran significativas, se procedió a construir un modelo con las variables significativas.

A este último modelo, se le volvieron a revisar los supuestos correspondientes. En el caso de valores extremos y valores de influencia sobre todos los valores ajustados no se observaron ninguno. Por ello, se siguió el análisis, se observó la prueba de homocedasticidad por medio de Breusch Pagan donde se obtuvo una probabilidad baja ($p < 0.001$), rechazando la hipótesis nula de que la variancia condicional de la variable respuesta es igual para cada valor de los predictores, con esto se asume que no hay homocedasticidad. Entonces como anteriormente ya se había aplicado una transformación a la respuesta, ahora se procedió a utilizar el método de mínimos cuadrados ponderados.

Este método obtiene la función de variancia o desviación estándar, donde luego se obtienen los pesos que son inversos a esa función y se aplican a los datos. Primeramente, se construyeron gráficos de los residuales absolutos de cada variable para observar patrones o forma de megáfono y como se observa en la Figura 9 (anexo) ambas variables presentaron una forma de megáfono. Con esto, se procedió a realizar el nuevo modelo que tiene el método de mínimos cuadrados ponderados y ya con ese modelo se analizaron otros supuestos.

El modelo no presentó valores extremos ni valores de influencia, además mantenía la normalidad. Con respecto a la linealidad, se realizó los gráficos de residuales parciales con respecto a cada variable predictora, y como se observa en la Figura 10 (anexo) en las variables numéricas se observa linealidad. Es por esto, que se procede a analizar la multicolinealidad del modelo.

En el análisis de la multicolinealidad, se procedió nuevamente a calcular el Factor de Inflación de la Variancia (VIF) y se demostró que ninguna de las variancias de los coeficientes de las variables predictoras sobrepasa a 10, asumiendo que el modelo no posee multicolinealidad.

A partir de aquí se analizó la significancia del modelo, y ambas variables quedaron, quedando como final el siguiente modelo

$$\mu_{Y|X} = \beta_0 + \beta_2 X_2 + \beta_5 X_5$$

Donde;

X_2 = Peso

X_5 = Comida

Con este modelo final, se calculó el $R^2_{ponderado}$, donde se observó que las variables peso, fuma, comida y tratamiento explican un 31% de la variabilidad de la respuesta.

A partir de eso, se procedió a determinar la ecuación de regresión del modelo.

$$\mu_{\log \log (y) | \text{Peso, Comida}} = 4.1608 + 0.0047\text{Peso} + 0.0094\text{Comida}$$

Por otra parte, para dar una interpretación con un aumento significativo, se debe multiplicar cada coeficiente en modo exponencial por su desviación estándar. Para el caso de peso, al aumentar el peso en 15 kg la presión arterial promedio aumenta 1.07 veces manteniendo las demás variables constantes, mientras que un aumento en el consumo de comida procesada diario de 3 comidas, la presión arterial promedio aumenta 1.03 veces, manteniendo las demás variables constantes. Es importante destacar que estos aumentos son muchas veces muy pequeños.

CONCLUSIONES

El modelo elegido contuvo únicamente dos variables que lograron explicar un 31% de la variabilidad ponderada de la Presión arterial, se podría decir que está en un límite puesto que la explicación no es mínima pero tampoco es moderada, pero está acorde con las variables que quedaron en el modelo. Dentro de los supuestos este modelo sigue conteniendo heterocedasticidad, pero utilizando el método de mínimos cuadrados ponderados se generó un modelo viable. Además de eso, el coeficiente de variación para el primer modelo propuesto fue de 35% de explicación, observando que por medio de este coeficiente es el primer modelo quien explica más, sin embargo, comparando los coeficientes de variación ajustados es el modelo final quien presenta un valor mayor. Por consiguiente, desde esta perspectiva es viable el uso del modelo final propuesto.

Las variables seleccionadas fueron peso y comida. Dentro de estas variables como lo explica Huerta, el aumento del peso puede aumentar los niveles de presión arterial (2001, p. 209), entonces puede tener sentido que fuese seleccionada, así como la variable comida, pues el consumo de comida procesada puede aumentar la cantidad de grasa en las arterias, que puede influir en el aumento de la presión arterial (Álvarez, Ramírez-Campillo, Martínez-Salazar, Vallejos-Rojas, Jaramillo-Gallardo, Salas Bravo, Cano-Montoya & Celis-Morales, 2016).

Cabe mencionar el hecho de que variables como la raza, edad, sexo no salieran como variables significativas puesto como lo explica (Huerta, 2001, p. 209), donde menciona que el sexo y la edad son variables importantes, menciona que en edades avanzadas las personas puedes ser más propensos a padecer hipertensión arterial, por otro lado, la raza como lo explica (González, 1999 & Balaguer, 1998, citados por Huertas, 2011, p.209) la raza tiene una mayor incidencia en los riesgos de padecer hipertensión arterial.

Con respecto a lo anterior, los expertos recomiendan para futuras investigaciones que haya una recolección de datos de personas con edades mayores, así como más variabilidad en las personas según raza puesto que podría ser determinante al momento de la construcción del modelo.

BIBLIOGRAFÍA

- Álvarez, C., Ramírez-Campillo, R., Martínez-Salazar, C., Vallejos-Rojas, A., Jaramillo-Gallardo, J., Salas Bravo, C., Cano-Montoya, J., y Celis-Morales, C. (2016). Hipertensión en relación con estado nutricional, actividad física y etnicidad en niños chilenos entre 6 y 13 años de edad. *Nutrición Hospitalaria*, 33(2), 220-225.
- Calvo, C. (2010). *Hipertensión arterial*. (2 ed.). Editorial Alfil. <https://elibro.proxyucr.elogim.com/es/lc/sibdi/titulos/40565>
- Espinosa-Brito, A. (2012). Mortalidad por hipertensión arterial. Tras la huella del “asesino silente”. *Revista Finlay*. <https://revfinlay.sld.cu/index.php/finlay/article/view/100>
- Fox, J. Weisberg S (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- Guevara, A., Ranero, V. M., Álvarez, V., Álvarez, E. N., y Tápanes, I. (2022). La hipertensión arterial como causa de muerte: causa básica vs causas múltiples. *Cuba*, 2013-2019. *Revista Habanera de Ciencias Médicas*, 21(4).
- Huerta, R. (2001). Factores de riesgo para la hipertensión arterial. *Arch Cardiol Mex*. 71(Suppl: 1), 208-210.
- Mejía-Rodríguez, O., Paniagua-Sierra, R., Del Refugio Valencia-Ortiz, M., Ruiz-García, J., Figueroa-Núñez, B y Roa-Sánchez, V. Factores relacionados con el descontrol de la presión arterial. *Salud Pública de México*. 2009;51(4):291-297. doi:10.1590/s0036-36342009000400005
- Rcore Team. (2022). *R: A Language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Sarkar, Deepayan. (2008). *Lattice: Multivariate Data Visualization with R*. Springer, New York. ISBN, 978-0-387-75968-5, <http://lmdvr.r-forge.r-project.org>.
- Universidad de Costa Rica. (2024). Suplemento OCI 2024: UCR en cifras. <https://www.ucr.ac.cr/medios/documentos/2024/suplemento-oci-2024-ucr-en-cifras-667dc33abe18e.pdf>

ANEXO

Cuadro. 2

Cuestionario para recolección de datos

Pregunta	Respuesta
¿Cuál es su edad en años cumplidos?	
¿Cuántas veces fuma usted al día? (Sea cigarro, vape, marihuana)	
¿Cuántas veces a la semana usted consume algún tipo de comida procesada? (incluye alimentos altos en azúcar, grasa o sal)	
¿Con cuál raza usted se identifica?	Marque con una x la seleccionada Blanca _____ Mestiza _____
¿Usted tiene algún familiar cercano que padezca hipertensión arterial?	Marque con una x la seleccionada Sí _____ No _____
¿Cuántas veces a la semana usted consume alcohol?	
¿Con cuál sexo usted se identifica?	Marque con una x la seleccionada Hombre _____ Mujer _____
¿Consume usted algún tratamiento con efecto hipertensivo?	Marque con una x la seleccionada Sí _____ No _____

Figura 3.

Diagrama de flujo para análisis de los datos

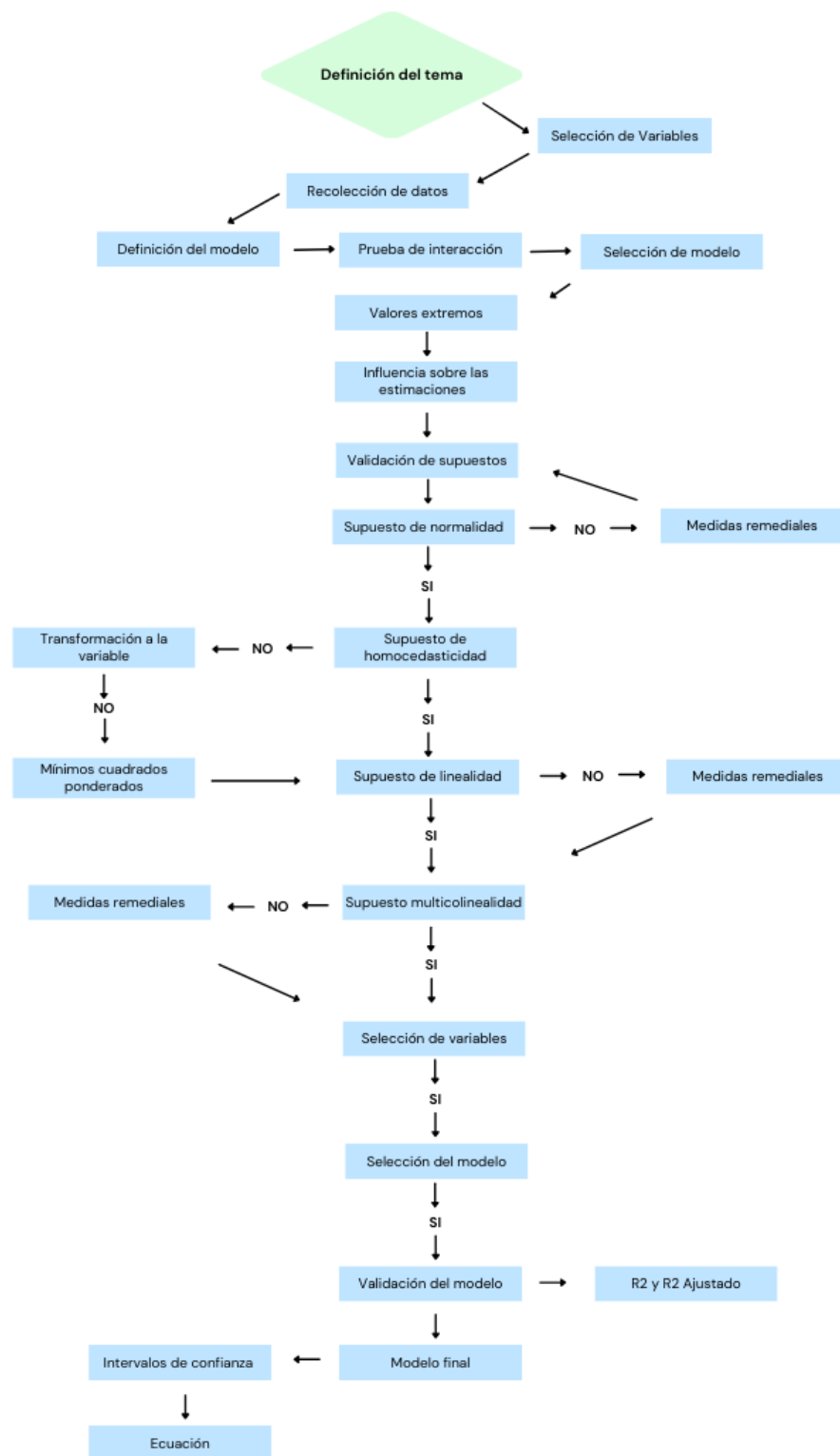


Figura 4.

Análisis gráfico de la relación entre la respuesta promedio y las variables predictoras.

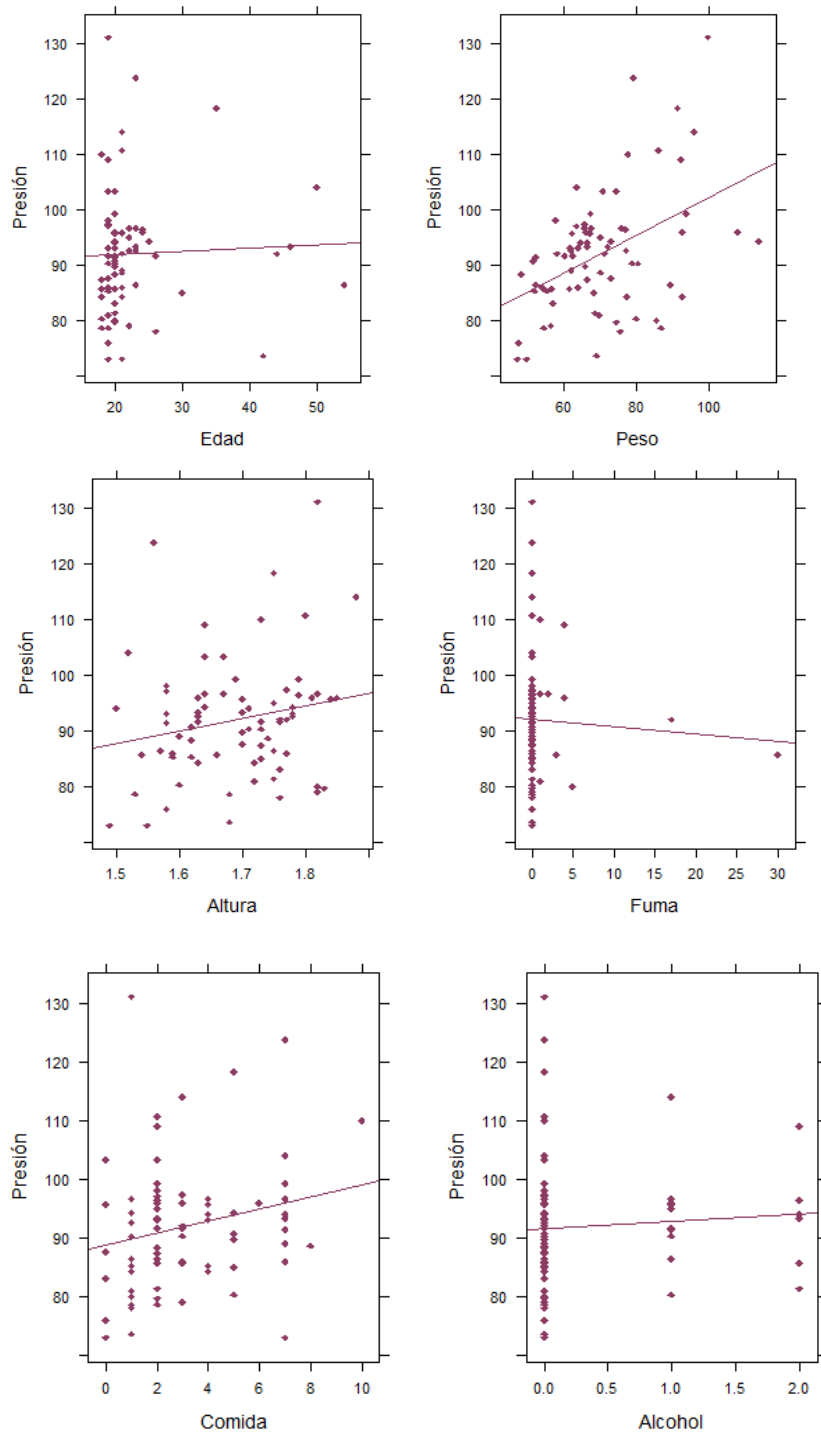


Figura 5

Interacciones de las variables categóricas con las variables predictoras

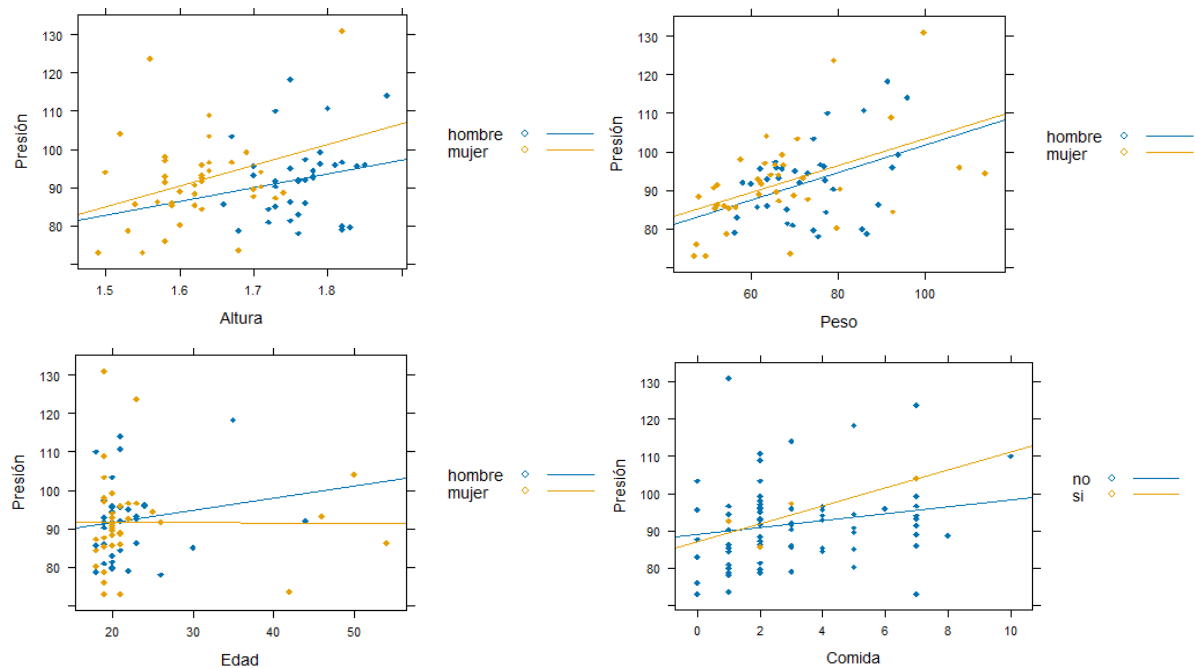


Figura 6

Residuales contra los valores ajustados

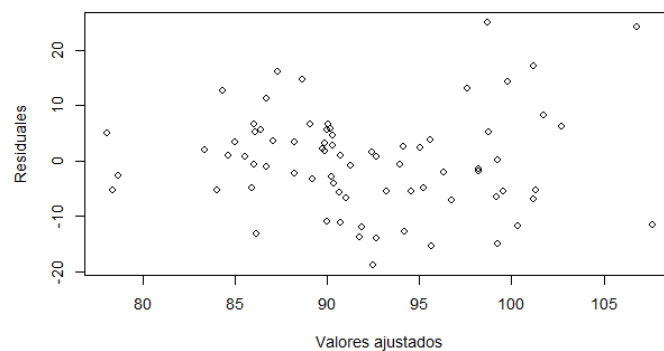


Figura 7

Transformaciones que maximizan la verosimilitud de la respuesta

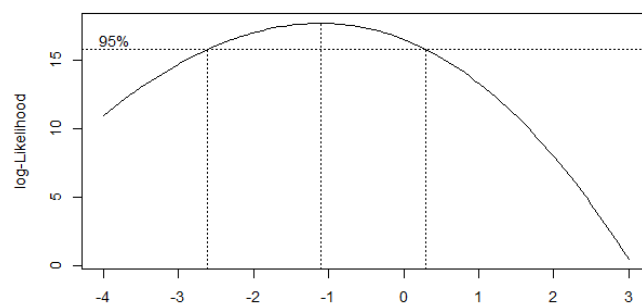


Figura 8

Residuales parciales contra las variables predictoras

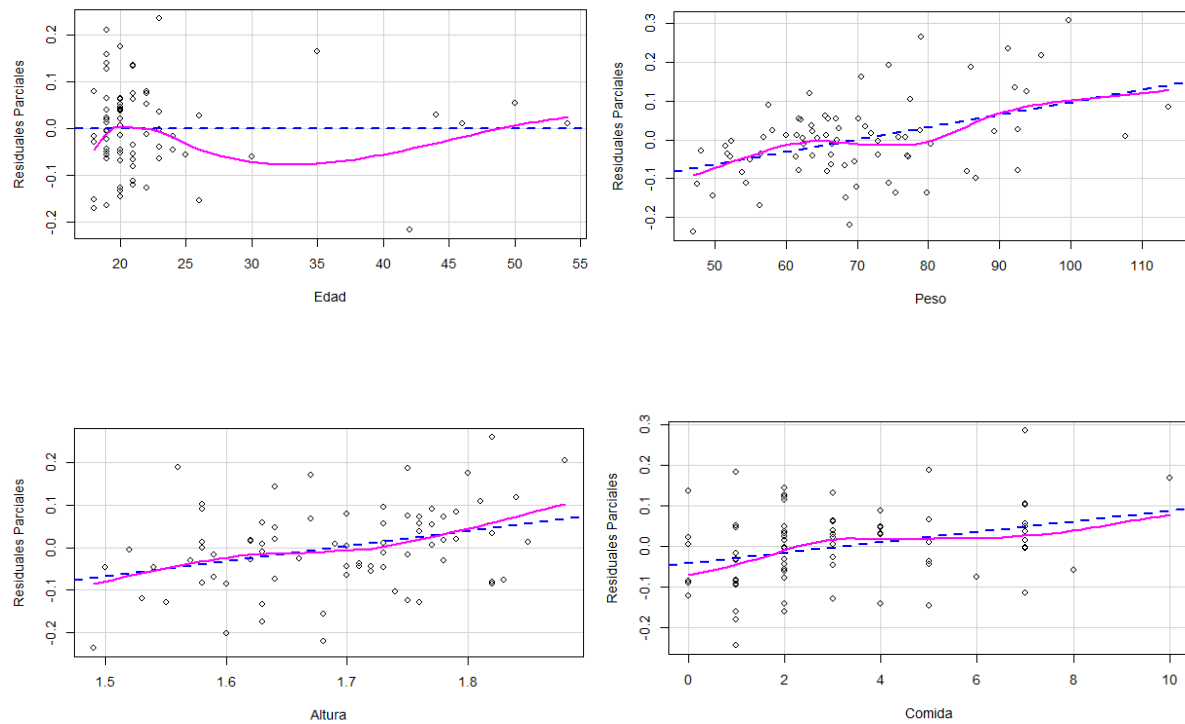


Figura 9

Residuales absolutos contra cada variable predictora

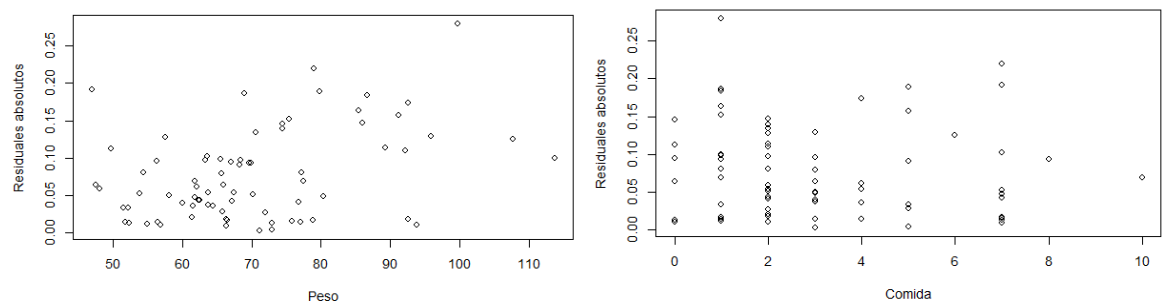


Figura 10

Residuales parciales contra las variables predictoras

