**Gonzalo Uriel Quintanilla Villanueva (20185271)**

**¿Qué factores que influyen en la percepción ciudadana respecto a la delincuencia en la ciudad de Lima Metropolitana?**

**Perfil de Github:** [GonzaloQV/Trabajo-Estad-stica-2-2021-2: Trabajo número 2 para el curso de Estadística para el análisis político (github.com)](https://github.com/GonzaloQV/Trabajo-Estad-stica-2-2021-2)

**Hipótesis:**

La pregunta a resolver en cuestión es ¿qué factores influyen en la percepción de delincuencia en Lima capital? Para ello, se esboza una hipótesis que plantea que ciertos factores pueden afectar la variable dependiente de percepción de delincuencia, los cuales mencionaré a continuación. Los factores a tomar en cuenta, y que servirían como variables independientes para el estudio, serían las variables de sexo, edad, estrato sociodemográfico, confianza institucional, percepción de seguridad en lugares públicos y privados, y la vigilancia policial. En primera instancia, se comprobará que estos factores sí pueden afectar en la sensación de la delincuencia a partir de pruebas estadísticas realizadas y revisión de literatura sobre el caso; en la introducción, se detallará más el proceso y caso del estudio.

**Introducción:**

La percepción de incremento o disminución de la delincuencia se puede definirse a partir de la inseguridad ciudadana que es “la perturbación angustiosa del ánimo que se deriva de la diferencia entre el riesgo percibido de ser víctima de un crimen y la victimización de hecho” (Aguilar et al, 2019). Esta sensación de inseguridad, puede deberse a diversos factores como en la ciudad que reside, el sexo de la persona, la edad, la confianza en sus autoridades, etc. Además, dicha inseguridad de los ciudadanos a ser víctima de algún tipo de delito suele ser latente dependiendo en qué entorno se desenvuelve y qué tanto hacen sus autoridades para resolver el problema. En el Perú, las entidades que deben velar por el bienestar de sus ciudadanos suelen ser el Gobierno y las municipalidades, sobre todo el primero, pues este tiene distintas herramientas para poder contrarrestar dicho flagelo. En el país, nueve de cada diez personas cree que será víctima de la delincuencia en los próximos 12 meses, según la Encuesta Nacional de Programas Estratégicos” (Kanashiro et al, 2020); esta sensación de inseguridad es preocupante; por ello, la acción que pueda emplear el Estado es relevante para solucionar el problema.

En este estudio, se hará énfasis en los diversos factores, algunos ya mencionados, que influyen en la percepción de inseguridad de los ciudadanos de la ciudad de Lima metropolitana, pues “Lima concentra más de la mitad de la población del Perú por lo que su importancia relativa en la mayoría de los indicadores entregados por las instituciones de justicia criminal es, al menos, similar” (Dammert 2012). Para ello, se tomó como referencia la base de datos situada en la página del INEI que titula “Encuesta nacional especializada sobre victimización”, el cual recoge datos de diversas partes del país para obtener la opinión de los peruanos respecto a la inseguridad ciudadana; por ello, esta se tomará como una variable independiente para explicar la variable dependiente, sobre todo la (in)seguridad en lugares como el centro comercial, el transporte público, el hogar, los mercados, los parques, la calle, etc. Este, entonces, se consideraría como un índice aditivo al agrupar varias variables y transformarla en la ya mencionada. En el siguiente párrafo, mencionaré el porqué de la elección de las variables independientes para explicar la variable dependiente.

Por un lado, se encuentra la variable independiente de Sexo, el cual contiene dos categorías Hombre y Mujer (exceptuando el género); en este caso se elige dicha variable pues puede explicar qué tan significativa podría ser, ya que como menciona Aguilar, lo que se esperaría es que las personas de sexo femenino se sientan físicamente más vulnerables frente a la acción delictiva; sin embargo, el sexo no es una variable determinante en la percepción de inseguridad en los habitantes (2019: 19); algo que por supuesto con este estudio se intentará comprobar. Por otro lado, se encuentra la variable de estrato sociodemográfico incluye la descripción de las características sociales y demográficas tales como nivel socioeconómico, lugar de residencia, ingresos, etc. Dicha variable podría determinar en qué medida los que tienen las categorías de los estratos, desde alta a baja, influyen en la percepción de inseguridad.

Por último, podemos resaltar la variable de confianza institucional, esta también es un índice aditivo que se compone de la opinión de los encuestados respecto a la confianza en distintas instituciones como la Policía Nacional, el Serenazgo, el Poder Judicial, el INPE, etc. Dicha variable nos puede ayudar a predecir si llega a influir o no en la variable dependiente. Cabe resaltar que el marco constitucional peruano declara en el artículo 166° que “La Policía Nacional tiene por finalidad fundamental garantizar, mantener y restablecer el orden interno. Presta protección y ayuda a las personas y a la comunidad. Garantiza el cumplimiento de las leyes y la seguridad del patrimonio público y del privado. Previene, investiga y combate la delincuencia. Vigila y controla las fronteras”.

**Metodología y Hallazgos**

La base de datos (data1) proporcionada por el INEI muestra los distintos resultados de las personas encuestadas en diversas partes del Perú. En primer lugar, para poder predecir la variable dependiente de percepción ciudadana sobre la delincuencia es importante realizar un ordenamiento y limpieza de la base de datos, con el fin de realizar idóneamente el estudio. Para ello, se debe escoger las variables independientes que logren predecir la variable dependiente; en este caso, como se mencionó, se escogieron las variables de Sexo, Edad, Estrato Sociodemográfico, Confianza Institucional, Percepción de seguridad en Lugares Públicos y en lugares privados, y vigilancia policía-serenazgo. Las tres primeras variables vienen predeterminadas en la base de datos, con lo cual no necesitan mayor modificación; sin embargo, en el caso de las últimas cuatro variables, estas en un principio se mostraban como variables categóricas con lo cual se cambiaron a numéricas. A continuación, describiré el cambio que se realizó en cada variable.

La primera variable convertida es la variable de Confianza, esta se creó realizando un índice aditivo de las variables de Confianza Institucional como confianza en la Policía, Serenazgo, el Poder Judicial, el Ministerio Público y el INPE; dichas variables al estar numeradas del 1 al 5, siendo 1 poca confianza y 5 mucha confianza, se suman y promedian todas para dar forma a la nueva variable de confianza. La segunda variable percepción de seguridad en lugares públicos, el cual también es un índice aditivo que sumó lugares como el Trabajo, la calle, el colegio, el mercado, el Centro Comercial, el banco, el parque y el transporte público; mientras que en la variable de percepción de seguridad en lugares privados se consideró lugares como la casa, el cajero automático y el vehículo propio. Por último, se encuentra la variable de Vigilancia, la cual suma las variables de percepción de vigilancia policial y de vigilancia del serenazgo en la zona de cada encuestado. Las variables que no fueron modificadas (Edad, Sexo y Estrato Sociodemográfico) solo fueron renombradas; es decir, en la variable Sexo solo mostraba valores de 1 y 2, 1 para Hombre y 2 para Mujer. Mientras que en la variable Sociodemográfica que era ordinal al ir del 1 al 3 fue renombrada en 1 como Estrato Bajo, 2 como Estrato Medio y 3 como Estrato Alta; y a la variable Edad se la dejó tal cual.

Una vez realizado el convertimiento de las variables categóricas a numéricas; se prosiguió a crear una base de datos (s0) en la que solo se escoja al departamento de Lima, específicamente solo Lima Metropolitana y los distritos que esta conforma; a su vez, se creo una nueva base de datos (data) en la que solo estén las variables idóneas para el estudio y que solo salgan los resultados de los encuestados en Lima metropolitana. Ahora bien, realizados este ordenamiento y limpieza de la base de datos se prosigue a hacer el modelo respectivo de Regresión Logística Ordinal, se escoge este tipo de regresión, pues nuestra variable dependiente de ‘Delincuencia’ se categoriza como ordinal, a tener valores de 1, 2 y 3 siendo 1 disminución delincuencial, 2 mantenimiento de la delincuencia y 3 aumento de la delincuencia. Para la realización del modelo de la mencionada regresión, se escogen como variables independientes a Edad, Sexo, Estrato Sociodemográfico, Confianza, Percepción de Seguridad en Lugares Públicos, Percepción de Seguridad en Lugares Privados y Vigilancia policía-serenazgo. El p-value de cada variable independiente nos arroja un valor menor a 0.05, a excepción de la variable sexo, Edad y estrato sociodemográfico; respecto a la variable Sexo se comprueba lo manifestado anteriormente por Aguilar respecto a que el sexo no es determinante para la percepción de inseguridad ciudadana (o de delincuencia). La predicción del modelo, que se corrobra a través del Pseudo R2, nos da un valor de 14.40%; en otras palabras, las variables independientes solo predicen en ese porcentaje a la variable dependiente.

La segunda prueba es también una prueba de Regresión Logística Ordinal, pero esta vez las variables independientes son componentes principales dados a través de un Análisis Factorial Exploratorio. Por un lado, se hace dicho análisis con las variables de Edad, Confianza, Seguridad en lugares públicos, en lugares privados y Vigilancia; para hacer este análisis factorial solo se pueden seleccionar variables numéricas, he ahí la razón por la cual en un inicio se transformaron estas variables categóricas en numéricas. Ahora bien, regresando al procedimiento del análisis factorial por componentes principales, se creó una nueva base de datos (data2), en la cual solo se encontraban las variables independientes antes mencionadas. A partir de dicha nueva base de datos, se hace una prueba KMO, en la cual el Overall MSA nos arroja un valor que es mayor a 0.05, con lo cual se prosigue con el estudio sin problemas. Luego, el R nos recomiendo utilizar solo dos componentes principales; o sea, las cinco variables independientes se agruparán en dos componentes con tal de reducir el número de variables. El primer componente contiene las variables de Vigilancia y Confianza; y el segundo componente contiene las variables de Edad, Seguridad en Lugares Públicos y Seguridad en Lugares Privados.

Una vez hecho el análisis factorial y al tener los dos componentes que serán de utilidad para predecir la variable dependiente de percepción ciudadana respecto a la delincuencia, se hace una nueva regresión logística ordinal, únicamente con los dos componentes que actuarán como variables independientes. El p-value de ambas variables es menor a 0.05 por lo cual se afirma la validez de las variables. Finalmente, se observa la predictibilidad del modelo, a partir de su Pseudo R2, el cual nos da un valor de 13.18%; en otras palabras, ambas variables independientes predicen en dicho porcentaje a la variable dependiente

**Conclusiones**

En conclusión, se corrobora que, a pesar de la realización de dos pruebas de Regresión Logística Ordinal con distintos métodos, la predicción de las variables independientes respecto a la variable dependiente de ‘Percepción ciudadana sobre la delincuencia’ no llega a ser significativo, pues tanto en el primer modelo como en el segundo los valores mostrados son muy bajos; es decir, con las variables independientes seleccionadas no se puede predecir a cabalidad la percepción de aumento o disminución de la delincuencia en los ciudadanos de Lima Metropolitana. La principal limitación surgida en esta investigación fue la base de datos, pues no se encontraban muchas variables relevantes que pudieran servir como variables independientes para predecir la variable dependiente; la base de datos extraída del INEI se basa en una encuesta; es decir, la mayoría de preguntas a responder eran de sí o no, con lo cual daban variables dicotómicas, las cuales muchas de ellas era irrelevantes para el estudio; y se complicó aún más pues la única variable numérica en la base de datos era la variable edad, las demás variables numéricas que se crearon a partir de índices aditivos se hizo con el fin de realizar una técnica de reducción de dimensiones, en este caso el Análisis Factorial Exploratorio por componentes principales, pues este análisis se hace únicamente con variables numéricas. Por ello, es que la interrogante a responder planteada al inicio no llega a ser resuelta tras lo descrito anteriormente. La recomendación más importante que se puede realizar es que se incluyan más variables relevantes para predecir los factores que influyen en la variable dependiente y así poder formular políticas públicas con el fin de apalear este flagelo que acaece en la sociedad no solo limeña, sino peruana.

**Bibliografía:**

Aguilar, A. y otros (2019). Relación entre el sexo y el estrato sociodemográfico de los habitantes de los distritos de San Isidro y Cercado de Lima con la percepción de la inseguridad ciudadana. <https://revistas.upch.edu.pe/index.php/RPH/article/download/3631/4031>

Dammert, L. (2012). Seguridad ciudadana en el Perú: Las cifras del desconcierto. *Lima: Corporación Andina de Fomento*. <https://www.academia.edu/download/30855406/Dammert_Peru_Crimen.pdf>

Constitución política del Perú.  
CONSTITUCION POLÍTICA DEL PERÚ (www.gob.pe)

Kanashiro, Dammert, L., & Hernández, W. (2020) La percepción de inseguridad ciudadana: determinantes y narrativas. Lima: *CIES*. <https://www.cies.org.pe/sites/default/files/files/articulos/economiaysociedad/la_percepcion_de_inseguridad_ciudadana_determinantes_y_narrativas.pdf>