

Método XSTATIS para el estudio de las principales causas de muerte en Ecuador 2020

XSTATIS method for the study of the main causes of death in Ecuador 2020

Sandra García-Bustos, Omar Ruiz-Barzola, Miguel Morán
y William Tandazo

*Departamento de Matemáticas, Centro de Estudios e Investigaciones
Estadísticas de la Escuela Superior Politécnica del Litoral, Ecuador*

Resumen

Desde inicios del año 2020 se han venido realizando un sin número de publicaciones científicas acerca de la enfermedad Covid-19. El presente trabajo tiene como objetivo usar la metodología de Análisis Triádico Parcial (PTA o XSTATIS) para estudiar las principales causas de muerte en Ecuador en 2020, año en que la Covid-19 estuvo presente, y cómo variaron en relación a los años previos con un enfoque particular en el efecto generado en los factores edad y género. El método XSTATIS estudia las posibles relaciones de k tablas bidimensionales, entonces para este estudio se definieron como dimensiones de cada tabla: las causas de muerte y edad en años cumplidos, considerándose tres tablas determinadas por el año: del 2018 al 2020. Esto se hizo por género es decir se aplicó XSTATIS para hombres y mujeres. Entonces, se logró identificar la interacción entre las comorbilidades registradas que afectaron a personas de determinada edad y género en los diferentes años estudiados, lo que permitió desarrollar un indicador a través de variables latentes, el cual muestra claramente el grado de riesgo de muerte, otorgándole una puntuación alta a las personas con mayor riesgo y puntuaciones más bajas a los de menor riesgo. Adicionalmente, los resultados mostraron que el 2020 fue un año atípico, pues presentó un incremento de las muertes de edades comprendidas entre 5 a 45 años, teniendo como causa primaria la Covid-19.

Palabras clave: Covid-19, Análisis triádico parcial PTA, X-STATIS, Indicador de riesgo.

Abstract

Since the beginning of 2020, countless scientific publications have been published about the Covid-19 disease. The objective of this work is to use the Partial Triadic Analysis (XSTATIS) methodology to study the main causes of death in Ecuador in 2020, the year in which Covid-19 was present, and how they varied in relation to previous years with a particular focus on the effect generated by age and gender factors. The By gender, that is, XSTATIS was applied for men and women. Then, it was possible to identify the interaction between the registered comorbidities that affected people of a certain age and gender in the different years studied, which allowed the development of an indicator through latent variables, which clearly shows the degree of risk of death, giving a high score for people at higher risk and lower scores for those at lower risk. Additionally, the results showed that 2020 was an atypical year as it presented an increase in deaths between the ages of 5 and 45, with Covid-19 as the primary cause.

Keywords: Covid-19, partial triadic analysis PTA, X-STATIS, risk indicator.

INTRODUCCIÓN

La pandemia de la Covid-19 a partir del año 2020 ha causado conmoción en la salud pública afectando alrededor de 230 países (WHO, 2022), colapsando sus sistemas sanitarios y causando afectación de índole económica-social. En Ecuador se reportó el primer caso el 29 de febrero de 2020, situación que se complicó a las pocas semanas después, provocando el confinamiento total desde el 17 de marzo del mismo año. Las cifras de fallecidos llegaron a sus picos más altos los días del 22 de marzo al 9 de junio de ese mismo. De acuerdo con Cevallos-Valdiviezo *et al.* (2021) los reportes generales descritos por el Ministerio de Salud Pública del Ecuador (MSP) no contemplaban o ajustaban al verdadero impacto que generó esta pandemia desde los meses de febrero a octubre 2020.

El análisis de datos se convirtió en una herramienta más para combatir el virus, pues su manejo e integración con metadatos previnieron en ciudades como Taiwán la saturación de hospitales (Chen *et al.*, 2021), manejando de manera más apropiada los servicios médicos durante los días más fuertes del brote. Además, mediante modelos estadísticos se logró determinar el efecto del sexo, edad y comorbilidades en personas infectadas (Li *et al.*, 2020; Zhang *et al.*, 2020; Ruiz Cantero, 2021).

En Ecuador, según el Instituto Nacional de Estadísticas y Censos (INEC) se registró en 2020 un exceso de 41 mil muertes con respecto al 2019 (INEC, 2021). De este excedente más de 50 por ciento de las defunciones fueron causadas por la Covid-19, siendo el resto como efecto de enfermedades respiratorias, cardíacas y diabetes. Medir el impacto que tuvo el año 2020 en distintos campos de estudio es fundamental para todo país que desee conocer con precisión la magnitud de este suceso en la población.

El método XSTATIS (*X Structuration des Tableaux a Trois Indices de la Statistique*) analiza la estructura de correlación entre diversas tablas y dentro de cada tabla mediante el uso de variables latentes (combinación lineal de variables estudiadas), obtenidas a través la descomposición de valores singulares de una matriz compromiso. Esta matriz compromiso resulta de combinar linealmente las tablas originales. En las secciones 2 y 3 se explica más sobre este método. El presente trabajo tiene como objetivo utilizar esta metodología para analizar las principales causas de muerte en Ecuador en 2020 comparándolo con los periodos 2018 y 2019, además de construir indicadores que muestren por género qué grupo de edades son

más vulnerables a las diversas causas de muerte de este estudio. Entonces, tenemos tres tablas, cada una definida por el año (2018, 2019, 2020), esto para hombres y mujeres. Cada tabla tiene como filas y columnas dos factores en este caso las filas son la edad y las columnas las causas de muerte y cada celda de cada tabla tiene los conteos de esas muertes. Entonces, como la descomposición de valores singulares, permite obtener proyecciones tanto de las filas y columnas en dimensiones de máxima variabilidad, es posible entender las relaciones causa de muerte y edad por género y construir indicadores a través de variables latentes obtenidas que muestren las subpoblaciones más vulnerables a las diversas causas de muerte. Adicionalmente, a través de los eigenvectores de la descomposición de valores singulares de la matriz compromiso, es posible proyectar cada tabla (2018, 2019 y 2020) para analizar cómo fue la estructura de muertes en el año 2020 y cómo varió en relación a los dos años previos. Como premisa, se espera que haya un incremento en las muertes en el año 2020 con respecto a los años previos en todas las edades, pero con mayor énfasis en las edades más jóvenes.

ANTECEDENTES

Existen varios trabajos sobre la evolución e impacto de la Covid-19 dentro del territorio ecuatoriano. Estudios realizados como el de Cuéllar *et al.* (2021) indican que hubo una correlación positiva entre la movilidad humana y los excesos de muertes en distintas provincias del Ecuador durante los primeros seis meses de 2020.

Otros investigadores han desarrollado modelos para estimar contagios como Fernández-Naranjo *et al.* (2021) quienes construyeron un modelo bayesiano para predecir la tendencia de contagios durante la pandemia. Por otro lado, otros también investigaron sobre correlación entre fallecer por Covid-19 y otros padecimientos (García-Bustos *et al.*, 2021), además del impacto a la salud mental y emocional de habitantes del Ecuador (Puchai-cela *et al.*, 2022). Otros trabajos del impacto de la Covid-19 en Ecuador son Carrión-Bósquez *et al.* (2022) y Zerna-Bravo *et al.* (2022).

Estos estudios dejan abiertas nuevas líneas de investigación al incrementar covariables relacionadas al desarrollo de la Covid-19 en Ecuador. Las metodologías de estadística multivariada buscan estudiar las varianzas y covarianzas entre variables, muchas de estas metodologías construyen variables ficticias para modelar las correlaciones entre variables (Castro, J. & Galindo, P., 2000). Una de estas metodologías, es el Análisis Triádico

Parcial (Thioulouse & Chessel, 1987; Kroonenberg, 1989; Thioulouse *et al.*, 2004; Rossi *et al.*, 2014).

El Análisis triádico Parcial (PTA) o XSTATIS, analiza arreglos tridimensionales; que para el presente estudio las dimensiones serían: la causa de muerte, el grupo de edad y la tercera los años; esto, debido a que se analizan los conteos de muertes por edades y por año. En este trabajo se estudian tres tablas simultáneas, una por año (2018-2020) para cada género, cada tabla estructurada en filas conformada por los grupos etarios y en columnas por las causas de muerte. Dado que el objetivo de este estudio es definir una estructura estable multivariante entre los diferentes años, es decir estudiar simultáneamente las submatrices para encontrar patrones comunes, el PTA es el método indicado para determinar la parte estable de las ocurrencias de las causas de muerte según la edad a través del tiempo y a su vez, medir cuánto difieren cada año en particular con respecto a esa estructura estable, para de esa forma determinar condiciones particulares de determinadas causas de muerte según determinado rango de edad en algún año en particular (González-Narváez *et al.*, 2020).

Existen otras metodologías de análisis multivariante de tres vías o con arreglos tridimensionales a las que pertenece el PTA denominada familia STATIS (Rodríguez, 2020); sin embargo, se aplica PTA debido a que es la única en la cual se mantienen estables las filas (grupos etarios) y columnas (causas de muerte) a través de la profundidad de la tercera dimensión (años). Las otras metodologías de análisis de tres vías requieren condiciones diferentes como, el no tener las mismas observaciones o individuos en las filas (STATIS Dual); el no medir las mismas características o variables en cada momento de la tercera dimensión (STATIS); contar solo con variables cualitativas (DISTATIS), entre otras (González-Narváez *et al.*, 2020).

El PTA usa la descomposición de valores singulares para poder representar la estructura de correlación entre tablas, filas y columnas de modo que construye variables latentes usando combinaciones lineales de las variables originales para representar las relaciones de filas, columnas y tablas (Medina, 2015), por lo que esas variables latentes servirán como indicadores en el presente trabajo.

Esta metodología ha sido utilizada en varios trabajos: Silmani *et al.* (2020) utilizaron PTA para la evaluación de la variabilidad espacio temporal de la calidad del agua en la cuenca del río Medjerda. Los autores analizaron 14 variables fisicoquímicas en 12 sitios de muestreo mensualmente durante 2013, es decir trabajaron con 12 tablas que eran los meses. Ellos destacaron que el análisis PTA les permitió identificar estructuras espacia-

les multivariantes comunes y problemas asociados con el mantenimiento de la calidad del agua, identificar los patrones consistentes en la química del agua, y permitir el análisis de la variabilidad temporal de la química del agua. Otros investigadores también han usado PTA para caracterizar variabilidad espaciotemporal en conjunto de datos ecológicos (Jiménez *et al.*, 2006; Rolland *et al.*, 2009; González-Narváez, 2021)

Varias investigaciones han hecho uso de variables latentes para construir indicadores como por ejemplo en el trabajo presentado por Vyas & Kumaranayake (2006) se utiliza análisis de componentes principales (PCA) para construir índices de estatus socioeconómico cuando hay ausencia de datos de ingresos y consumo. En ese trabajo se construyeron los indicadores usando PCA sobre la propiedad de activos duraderos, el acceso a servicios públicos e infraestructura, y las variables características de la vivienda. La principal ventaja de este método sobre los métodos más tradicionales basados en los ingresos y los gastos de consumo es que evita muchos problemas como el sesgo de recuerdo, la estacionalidad y el tiempo de recopilación de datos. Rygel *et al.* (2006) también propusieron una metodología para construir indicadores compuestos de vulnerabilidad en las que combinaban un análisis de Pareto y PCA.

DATOS Y METODOLOGÍA USADA EN LA INVESTIGACIÓN

Datos

Para este estudio se han considerado las bases de datos de las defunciones generales de Ecuador publicadas por el INEC (2021) para los años 2018, 2019 y 2020. El INEC utiliza el formulario estadístico de defunción como requisito para el registro de los fallecidos en las oficinas del registro civil. De esta herramienta se obtienen luego las bases de datos de defunciones generales dentro del territorio ecuatoriano.

Ecuador tiene cuatro regiones: costa, sierra, oriente e insular, y la mayor parte de sus regiones sierra y oriente están formadas por poblaciones rurales (Mera-Intriago, 2023). Además, Ecuador presenta subregistros en mortalidad, el cual es más persistente en las áreas rurales y en la pandemia ese problema se agravó (Vera *et al.*, 2019). A fin de minimizar el problema de los subregistros, solo se han considerado variables como sexo, edad, causa de muerte y año de defunción para el estudio de las causas de mortalidad. La variable edad ha sido categorizada considerando las edades mayores a cinco años, y formando grupos considerando un rango de cinco años, es decir de (5,10], (10,15], ..., (85 o más), en total 17 grupos de edades.

Mientras que para la variable causas de muerte, categórica también, se seleccionaron las 18 principales causas de muerte registradas en el informe anual de 2020 del INEC codificadas de acuerdo a la CIE-10 (WHO, 2019), la Décima Revisión de la Clasificación Internacional de Enfermedades (INEC, 2021b). De modo, que los valores dentro de cada tabla construida son de tipo conteo de muertes ocurridas por edad y causa. La Tabla 1 presenta el código de estas causas:

Tabla 1: Causas de muertes analizadas

Causas	Grupo de enfermedad
C16, C22, C34	Tumores Malignos neoplasias (T)
E11, E14	Diabetes mellitus (D)
I10, I11	Enfermedades hipertensivas (H)
J12, J15, J18	Influenza, gripe y neumonía (VR)
K74	Fibrosis y cirrosis del hígado (HP)
N18	Insuficiencia renal crónica (RN)
COVID-19	Covid-19(CV)
I21	Infarto agudo del miocardio (I)
I61, I63, I64, I67	Enfermedades cerebrovasculares (C)

Fuente: elaboración del autor a partir de la CIE-10

Análisis Triádico Parcial o XSTATIS

Esta metodología multivariante tiene como objetivo analizar el comportamiento de un grupo de variables medidas a través del tiempo en un conjunto de unidades de investigación (González-Narváez *et al.*, 2020). Las variables y las unidades de investigación son las mismas a través del tiempo (n unidades de investigación, p variables y k periodos analizados) (Slimani *et al.*, 2020), es decir, se estudian las interrelaciones entre k tablas X_i , $i = 1, \dots, k$ organizadas en n filas y p columnas

Para este trabajo, $n=17$ son los grupos de edades, $p = 18$ categorías para la variable causas de muerte y k es igual a tres años (2018, 2019, 2020).

Cada uno de los k periodos tiene los tripletes $(X_1, D_n, D_p), \dots, (X_k, D_n, D_p)$, siendo D_n la matriz diagonal con los pesos para las filas y D_p la matriz diagonal con los pesos para las columnas (González-Narváez *et al.*, 2020). A partir de esto, se desarrollan los tres pasos que requiere la metodología y que se describen a continuación.

El paso de la interestructura permite obtener los pesos para construir una tabla que resuma las k tablas estudiadas (Thioulouse, 2011). Los pesos son obtenidos a partir del primer eigenvector de la matriz RV o de correlaciones vectoriales entre las matrices. La matriz RV indica como están correlacionadas las tablas de datos. Por ejemplo, para el presente estudio esta matriz RV sería de 3×3 , y una correlación alta positiva indicaría que ambas tablas tienen una tendencia similar, es decir a mayores valores de muertes en una celda mayores valores en la misma celda en la otra tabla. La matriz RV o de correlaciones vectoriales entre cada combinación de par de tablas de datos (X_i, X_j) se la define como:

$$RV(X_i, X_j) = \frac{\text{traza}(X_i^t D_n X_j D_p)}{\sqrt{\text{traza}(X_i^t D_n X_i D_p)} \sqrt{\text{traza}(X_j^t D_n X_j D_p)}}$$

Para el presente trabajo cada matriz de datos X_i es de dimensión 17 (grupo de edades) por 18 (causas de muerte) y se tiene una matriz de datos por cada año; además, D_n es la matriz diagonal de pesos de las edades, es decir $Dn = \text{diag}(w_1, w_2, \dots, w_n)$ y $w_i = 1/n$; mientras que D_p es la matriz identidad de dimensión p correspondiente a los pesos para las causas de muerte.

El siguiente paso fue obtener la matriz compromiso a través de una combinación lineal de las k tablas, los coeficientes de la combinación lineal son obtenidos del primer eigenvector de la matriz RV. La matriz compromiso maximiza la media de las correlaciones entre las variables de cada matriz. X_i Proporciona una imagen de las estructuras que son comunes a todas las tablas (Medina, 2015).

Cabe mencionar aquí que la descomposición de valores singulares de una matriz M , utilizada en este método, está dada por la siguiente identidad, $M_{n \times p} = U \Sigma V^T$. En el caso de este estudio M es la matriz compromiso. Donde Σ es una matriz diagonal formada por las raíces cuadradas de los valores propios no nulos ordenados de mayor a menor de las matrices $M^T M$ y MM^T ; U y V son bases ortonormales formadas por los vectores propios correspondientes de la matrices MM^T y $M^T M$. Entonces a través de las combinaciones lineales tomando como coeficientes los vectores que corresponden a los mayores eigenvalores de U y V en las filas o columnas de la tabla se pueden construir proyecciones en esos espacios de máxima variabilidad.

Por último, el análisis intraestructural resume la variabilidad en las tablas en torno a la estructura común definida por el compromiso y en él se

destacan los elementos que mejor encajan (o no encajan) a la estructura del compromiso (Silmani *et al.*, 2020). Entonces, a través de una descomposición de valores singulares de la matriz compromiso y de cómo las proyecciones de las filas y columnas se visualizan el mapa de factores, se pueden interpretar las similitudes de las tablas.

Más detalle de este método se pueden encontrar en Abdi *et al.* (2012). El cálculo y la presentación gráfica fueron obtenidos utilizando el paquete ade4 (Thioulouse *et al.*, 2018) para el software R (R Core Team, 2022).

Con la finalidad de dar soporte a los resultados obtenidos con el PTA, se aplicó la versión multivariante no paramétrica del análisis de varianza, conocida como PERMANOVA (*Permutational multivariate analysis of variance*) propuesto por Anderson, M. (2001 y 2017), el cual es un método estadístico que sirve para comparar grupos de elementos con el fin de probar la hipótesis nula, la cual indica que los centroides y la dispersión son equivalentes entre todos los grupos. Basa su cálculo en la distancia -generalmente Bray-Curtis- entre dos objetos incluidos en el estudio.

Para el caso de un solo factor, escalable para múltiples factores, es el siguiente: Se obtiene la Suma Cuadrática Total, SS_T , calculando el cuadrado de las distancias entre todas las observaciones, dividido para N el número total de observaciones. Luego se obtiene la Suma Cuadrática dentro de grupos, SS_W , calculando el cuadrado de las distancias entre las observaciones de un mismo grupo, dividido para n el tamaño del grupo. Seguidamente se obtiene por diferencia la Suma Cuadrática entre los m grupos, $SS_A = SS_T - SS_W$. Con ello se obtiene el pseudo estadístico $F = [SS_A / (m-1)] / [SS_W / (N-m)]$. Finalmente se obtiene el valor- $p = [\text{número de permutaciones con } F^* > F] / [\text{Total de permutaciones}]$, siendo F^* un nuevo valor de F obtenido en cada permutación; y con este valor se toma la decisión de rechazar o no la hipótesis nula; por tal motivo, el número de permutaciones es importante, esto brinda mayor solidez al resultado obtenido.

Para nuestro caso en particular se probó el modelo multifactorial multinivel con interacciones, considerando como factores de interés o como efectos principales la edad en rangos, el género y los años. El cálculo del valor - p se realizó con 999 permutaciones (valor sugerido por el autor), y fue obtenido utilizando el paquete vegan (Oksanen *et al.*, 2022) versión 2.6-4, para el software R (R Core Team, 2022); se consideró un nivel de significación de cinco por ciento.

RESULTADOS

Estadística descriptiva de las causas de muerte por género y edad

El presente trabajo considera 56,317 muertes ocurridas en el año 2020, 27,025 para mujeres y 29,292 para hombres, pues se excluyen los casos de niños menores de cinco años para ambos sexos y solo se está tomando en cuenta las causas mencionadas en la Tabla 1. Para mayor facilidad se van a usar para esta sección los códigos T, D, H, VR, HP, RN, CV, I y C, tal como está en la Tabla 1.

La Tabla 2 presenta una comparación porcentual entre los totales de muertes por las causas analizadas en este trabajo para los años 2018, 2019 y 2020, tanto para hombres y mujeres, valores en negro y rojo respectivamente. Si consideramos los años 2019 y 2018, se puede apreciar incrementos en muertes de hombres y mujeres para las causas analizadas del grupo T (Tumores malignos), I (Infartos), RN (Insuficiencia renal crónica), esos incrementos son alrededor de diez por ciento. Si nos enfocamos en el 2020, año en que aparece la Covid-19 y lo comparamos con 2019, se puede apreciar que en las mujeres hubieron incrementos considerables en las causas relacionadas con Diabetes (D), Hipertensión (H), Infartos (I), Virales respiratorias (VR), e Insuficiencia renal crónica (RN) de 49, 59, 79, 31 y 31 por ciento, respectivamente. Mientras que, en los hombres, casi todas las causas presentaron reducción menos VR. Cabe mencionar que en la base de datos analizada se registraron 15,670 y 8,068 casos para hombres y mujeres de la Covid-19 (CV) respectivamente.

Tabla 2: Comparativo porcentual de los casos de muerte analizados

Año	T	D	H	I	C	VR	HP	RN
2019								
vs	8%	10%	-5%	8%	-2%	-2%	-1%	8%
2018	9%	-2%	1%	9%	1%	4%	-3%	10%
2020								
vs	-6%	7%	-27%	-11%	-43%	36%	-27%	-35%
2019	-7%	49%	59%	79%	7%	31%	-12%	31%

Tumores malignos (T), Diabetes mellitus (D), Hipertensión (H), Influenza, gripe y neumonía (VR), Covid-19 (CV), Fibrosis y cirrosis hepática (HP), Insuficiencia Renal (RN), Infarto Agudo de miocardio (I), Enfermedad cerebrovascular (C).

Fuente: elaboración del autor usando datos de Defunciones Generales, 2018-2020, INEC.

La Tabla 3 resume la distribución porcentual de los registros totales (2018-2020) de muerte por causa, edad y género.

Tabla 3: Distribución de muertes en Mujeres y hombres por grupo de edades 2018-2020

Edades	Mujeres								Covid-19
	T	D	H	I	C	VR	HP	RN	
(5,50]	15%	13%	3%	18%	10%	13%	4%	3%	21%
(50,60]	12%	20%	4%	17%	7%	8%	6%	4%	22%
más de 60	8%	15%	10%	26%	9%	12%	4%	3%	14%
	Hombres								Covid-19
	T	D	H	I	C	VR	HP	RN	
(5,45]	9%	7%	2%	26%	8%	17%	5%	3%	23%
(45,55]	8%	13%	3%	21%	7%	10%	7%	3%	29%
más de 55	8%	11%	7%	26%	8%	11%	4%	3%	22%

Tumores malignos (T), Diabetes mellitus (D), Hipertensión (H), Influenza, gripe y neumonía (VR), Covid-19 (CV), Fibrosis y cirrosis hepática (HP), Insuficiencia Renal (RN), Infarto Agudo de miocardio (I), Enfermedad cerebrovascular (C).

Fuente: elaboración del autor a partir de los datos de INEC.

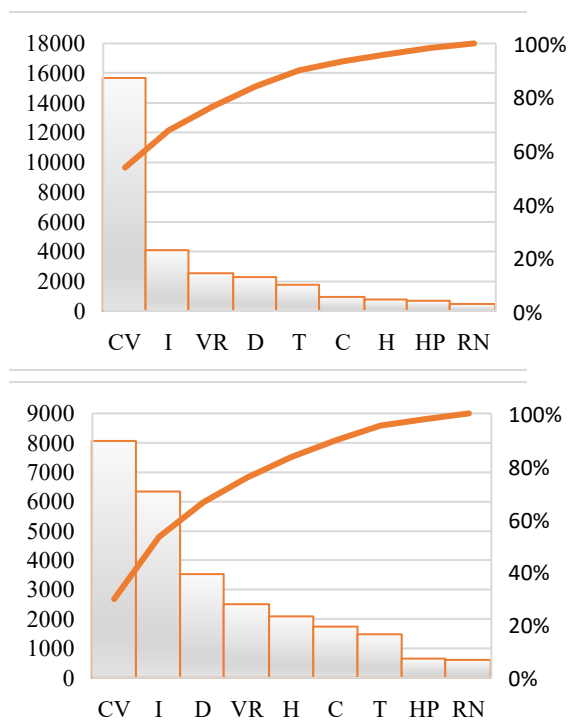
La Figura 1 presenta diagramas de Pareto para hombres y mujeres, en los cuales se puede ver los conteos de muerte en el año 2020 por cada una de las 18 causas consideradas en este estudio y que por grupos de enfermedad a los que pertenecen de acuerdo a la CIE-10 mencionados en la Tabla 1.

La Figura 1 presente dos ejes y, el primero con las frecuencias absolutas y el otro con las frecuencias relativas acumuladas. Las barras se rigen con la frecuencia absoluta y la línea naranja es una ojiva que grafica la frecuencia relativa acumulada. Por ejemplo, para el primer gráfico de esta figura (hombres), la causa CV presentó aproximadamente 16 mil defunciones para los hombres, es decir cerca de 60 por ciento de muertes por esta causa, además, aproximadamente 80 por ciento de las muertes fueron por las causas codificadas como CV, I VR, mientras que en las mujeres (segundo gráfico) se aprecia que cerca de 80 por ciento de las muertes analizadas fueron por las causas CV, I, D y VR.

La Figura 2 muestra las muertes por edades de las tres principales causas para hombres y mujeres.

Para ambos casos hombres y mujeres la principal causa de muerte registrada fue el Covid-19, para todas las edades, teniendo un incremento importante a partir de los 35 años.

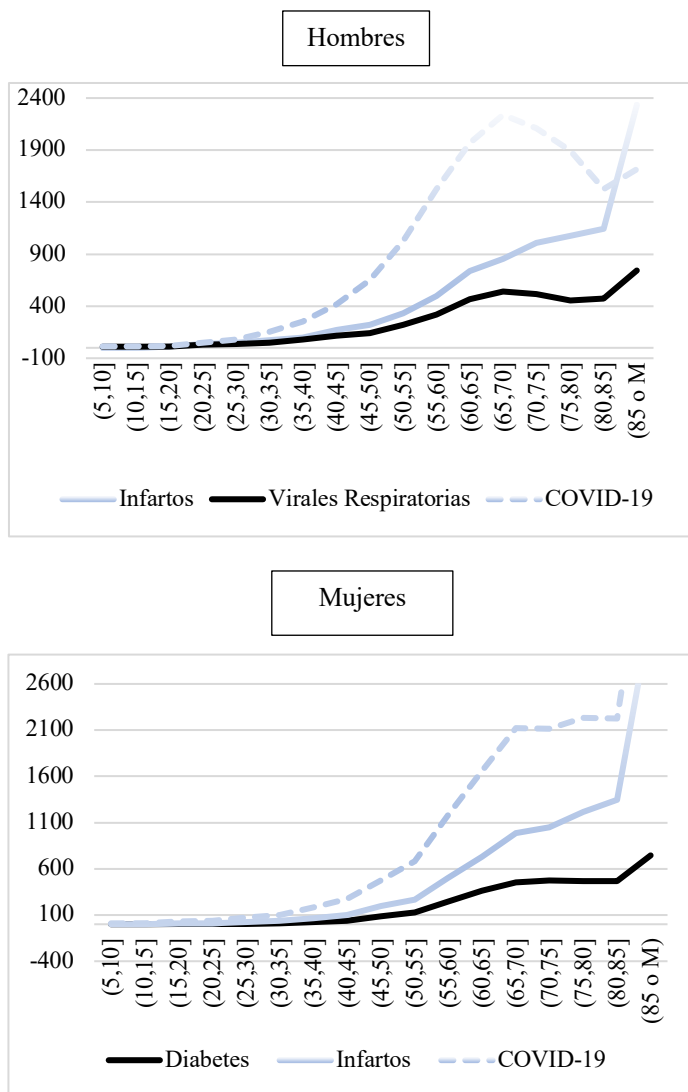
Figura 1: Diagrama de Pareto, muertes analizadas 2020, hombres y mujeres



Tumores malignos (T), Diabetes mellitus (D), Hipertensión (H), Influenza, gripe y neumonía (VR), Covid-19 (CV), Fibrosis y cirrosis hepática (HP), Insuficiencia Renal (RN), Infarto Agudo de miocardio (I), Enfermedad cerebrovascular (C).

Fuente: elaboración del autor usando datos de Defunciones Generales, 2018-2020, INEC.

Figura 2: Muertes por edades hombres y mujeres 2020



Fuente: elaboración del autor usando datos de Defunciones Generales, 2018-2020, INEC.

Método XSTATIS: Interestructura

La interestructura parte del cálculo de la matriz de correlaciones vectoriales RV entre los distintos años. En hombres, las muertes del 2018 tuvieron una correlación de 0.94 con las de 2019, mientras que las correlaciones entre las muertes de 2020 con las del año 2018 y 2019 fueron de 0.69, positiva pero no tan alta. Por otro lado, en las mujeres, las muertes del año 2018 tuvieron una correlación de 0.98 con las del año 2019, habiendo un descenso en las correlaciones entre las muertes del año 2020 con las de 2019 y 2018, siendo de 0.78.

A partir de la matriz RV se obtienen los eigenvalores y scores, en este caso la primera dimensión explica 90 y 85 por ciento de la variabilidad de las muertes ocurrida entre los años para hombres y mujeres, respectivamente. Mientras que la segunda dimensión explica 9.7 y 13 por ciento de la variabilidad entre los años para hombres y mujeres, respectivamente.

La Figura 3 muestra la representación de los scores en las dos dimensiones, se puede apreciar la alta correlación entre los años 2018 y 2019, mientras que 2020 se presenta diferente que los años 2018 y 2019.

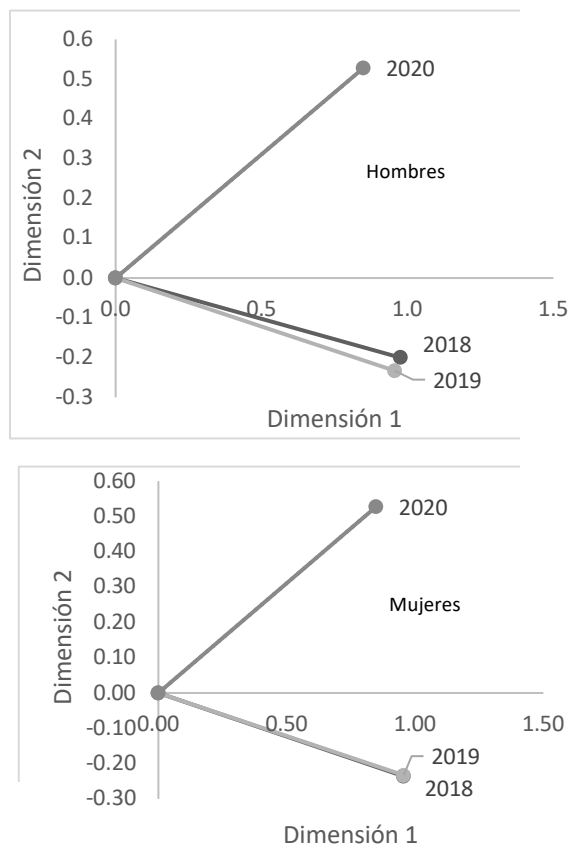
Cálculo de la matriz compromiso

La matriz compromiso se obtiene combinando linealmente las tablas de datos de los años 2018, 2019, 2020, como se mencionó anteriormente se lo obtiene mediante el primer eigenvector de la matriz RV. Para el caso de hombres, los coeficientes fueron 0.954, 0.955 y 0.852, mientras que para las mujeres fueron 0.974, 0.972 y 0.896 para los años 2018, 2019 y 2020, respectivamente. Es decir, en ambos casos se dio mayores ponderaciones a los años 2018 y 2019.

Al realizar la descomposición de valores singulares en la matriz compromiso se retuvieron dos dimensiones que explicaban 97.5 por ciento para hombres y 98.65 por ciento para mujeres. Los *scores* de los grupos de edad y de las variables se pueden visualizar en las figuras 4 y 5.

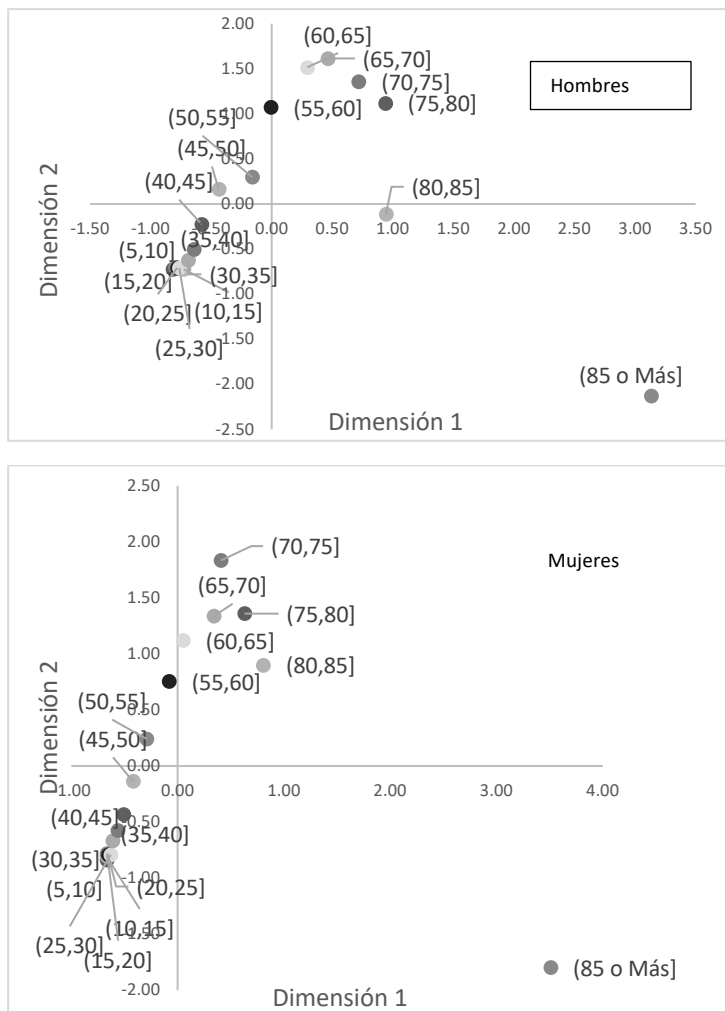
En la Figura 4, se puede destacar que desde los cinco años hasta antes de los 45 años están representados por puntuaciones negativas en la primera dimensión para hombres y en el caso de mujeres desde los cinco años hasta 50 años, es decir que en estos grupos es a donde se registró menos muertes en todas las causas (debido a que los coeficientes de las variables en la primera dimensión son positivos) y si consideramos las coordenadas negativa de la segunda dimensión las principales causas de muerte fueron las enfermedades Covid-19, Infartos (I), Influenza, Gripe y Neumonía (VR).

Figura 3: Interestructura para las defunciones de Hombres y Mujeres 2018-2019 vs 2020



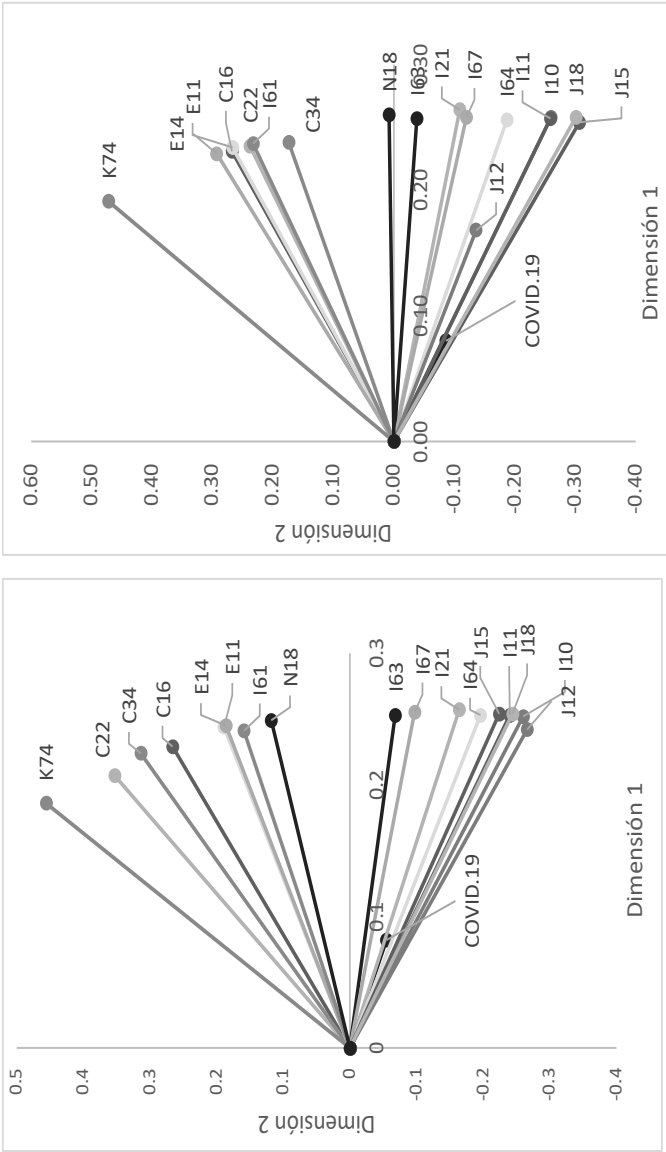
Fuente: elaboración del autor usando datos de Defunciones Generales, 2018-2020, INEC.

Figura 4: Grupos Etarios matriz compromiso hombres y mujeres



Fuente: elaboración del autor usando datos de Defunciones Generales, 2018-2020, INEC.

Figura 5: Scores causas de muerte para hombres y mujeres



Tumores malignos (C16, C22, C34), Diabetes mellitus (E11, E14), Hipertensión (I10, I11), Influenza, gripe y neumonía (J12, J15, J18), Covid-19, Fibrosis y cirrosis hepática (K74), Insuficiencia Renal (N18), Infarto Agudo de miocardio (I21), Enfermedad cerebrovascular (I61, I63, I64, I67)

Fuente: elaboración del autor usando datos de Defunciones Generales, 2018-2020, INEC.

Mientras que las edades de 45 a 55 años en hombres y de 50 a 60 años en mujeres tienen puntuaciones positivas en la segunda dimensión y negativas en la primera dimensión, es decir que en estos grupos ya empiezan a aparecer como causas principales de muerte, las enfermedades tumores malignos (T) y diabetes mellitus D (D).

Por otra parte, las edades mayores a 55 en hombres y a 60 en mujeres ya muestran puntuaciones positivas en ambas dimensiones, pudiendo morir por cualquier causa, en especial se presentan puntuaciones altas en la primera dimensión en las edades mayores a 85 años.

En la Figura 5 se puede visualizar que todas las causas analizadas en este estudio tienen cargas vectoriales positivas en la dimensión uno. En relación a la dimensión dos tienen cargas positivas las relacionadas a HP, T, D, RN y la causa I61, hemorragia intraencefálica que está dentro del grupo (C) de enfermedades cerebrovascular y cargas negativas las relacionadas a los grupos VR, CV, I, H, las restantes del grupo C. Las líneas más cercanas indican una mayor correlación entre sí. Por ejemplo, en hombres aparecen muy relacionadas entre sí Diabetes mellitus (E11, E14) y Tumores malignos (C16, C22, C34) y en mujeres, Diabetes mellitus, insuficiencia renal (N18). También se resalta la alta correlación entre la Covid-19 y los grupos de influenza, neumonía y gripe, infarto, hipertensión, lo que coincide con resultados de algunos estudios (Wu & McGoogan, 2020, García-Bustos *et al.*, 2021).

La Intraestructura

Finalmente, en la intraestructura se proyectan las filas y columnas de cada una de las k tablas iniciales en el espacio compromiso, lo cual está representado en las figuras 6 y 7.

Las figuras 6 y 7 presentan las puntuaciones factoriales para las diferentes edades considerando dos dimensiones y las cargas vectoriales de las causas para los años 2018-2020 y por género, hombre y mujer. Durante el 2020 se observa como todos los rangos de edad se están agrupando de forma muy distinta con respecto al 2018 y 2019. Además, como durante el año 2020 todas las causas de muerte se están agrupando en la parte inferior del plano, lo que indica la alta correlación entre Covid-19 y las demás causas de muerte afectando tanto a hombres y a mujeres.

The figure consists of two scatter plots side-by-side, both with 'Dimensión 1' on the x-axis and 'Dimensión 2' on the y-axis.

Left Plot: The x-axis ranges from -5.00 to 15.00, and the y-axis ranges from -4.00 to 3.00. It shows a dense cluster of data points, many labeled with coordinates and the year 2018. Notable points include (60,65).2018, (55,60).2018, (70,75).2018, (75,80).2018, (50,55).2018, (45,50).2018, (40,45).2018, (35,40).2018, (5,10).2018, (10,15).2018, (30,35).2018, (15,20).2018, (20,25).2018, (25,30).2018, (80,85).2018, and (85,0).2018.

Right Plot: The x-axis ranges from 0 to 1, and the y-axis ranges from -0.4 to 1. It shows a more spread-out distribution of data points, also labeled with coordinates and the year 2018. Notable points include K74.2018, E14.2018, C22.2018, E11.2018, C16.2018, C34.2018, I63.2018, N18.2018, I21.2018, I67.2018, I64.2018, I11.2018, I10.2018, J18.2018, and J15.2018.

The figure consists of two scatter plots side-by-side, both showing the relationship between Dimension 1 (X-axis) and Dimension 2 (Y-axis).

Left Plot: The X-axis (Dimension 1) ranges from 0.00 to 15.00, and the Y-axis (Dimension 2) ranges from -4.00 to 3.00. The data points are labeled with coordinates (Year, Year) and connected by a line, showing a clear downward trend from 2019 to 2018. The points are: (60,65).2019, (65,70).2019, (70,75).2019, (75,80).2019, (55,60).2019, (50,55).2019, (45,50).2019, (40,45).2019, (35,40).2019, (30,35).2019, (25,30).2019, (20,25).2019, (15,20).2019, (10,15).2019, (80,85).2019, and (85,0).2019.

Right Plot: The X-axis (Dimension 1) ranges from 0 to 1, and the Y-axis (Dimension 2) ranges from -0.4 to 1. The data points are labeled with years and connected by lines, showing a more complex, non-linear relationship. The points are: K74.2019, C16.2019, E11.2019, E14.2019, C34.2019, C22.2019, I61.2019, N18.2019, I63.2019, I21.2019, I67.2019, I64.2019, I10.2019, I11.2019, I15.2019, and I18.2019.

The figure consists of two PCA plots. The left plot shows the first two principal components (Dimensión 1 and Dimensión 2) with data points labeled by date ranges and case/death counts. The right plot shows the same data points with labels for specific dates and case/death counts.

Left Plot Data (Approximate Coordinates):

Label	Dimensión 1 (X)	Dimensión 2 (Y)
(75,80].2020	2.0	0.8
(80,85].2020	2.5	0.7
(70,75].2020	1.5	0.6
(65,70].2020	1.0	0.5
(60,65].2020	0.5	0.4
(55,60].2020	0.0	0.3
(50,55].2020	-0.5	0.2
(45,50].2020	-1.0	0.1
(40,45].2020	-1.5	0.0
(35,40].2020	-2.0	-0.1
(30,35].2020	-2.5	-0.2
(25,30].2020	-3.0	-0.3
(20,25].2020	-3.5	-0.4
(15,20].2020	-4.0	-0.5
(10,15].2020	-4.5	-0.6
(5,10].2020	-5.0	-0.7
Más o	16.0	-0.2

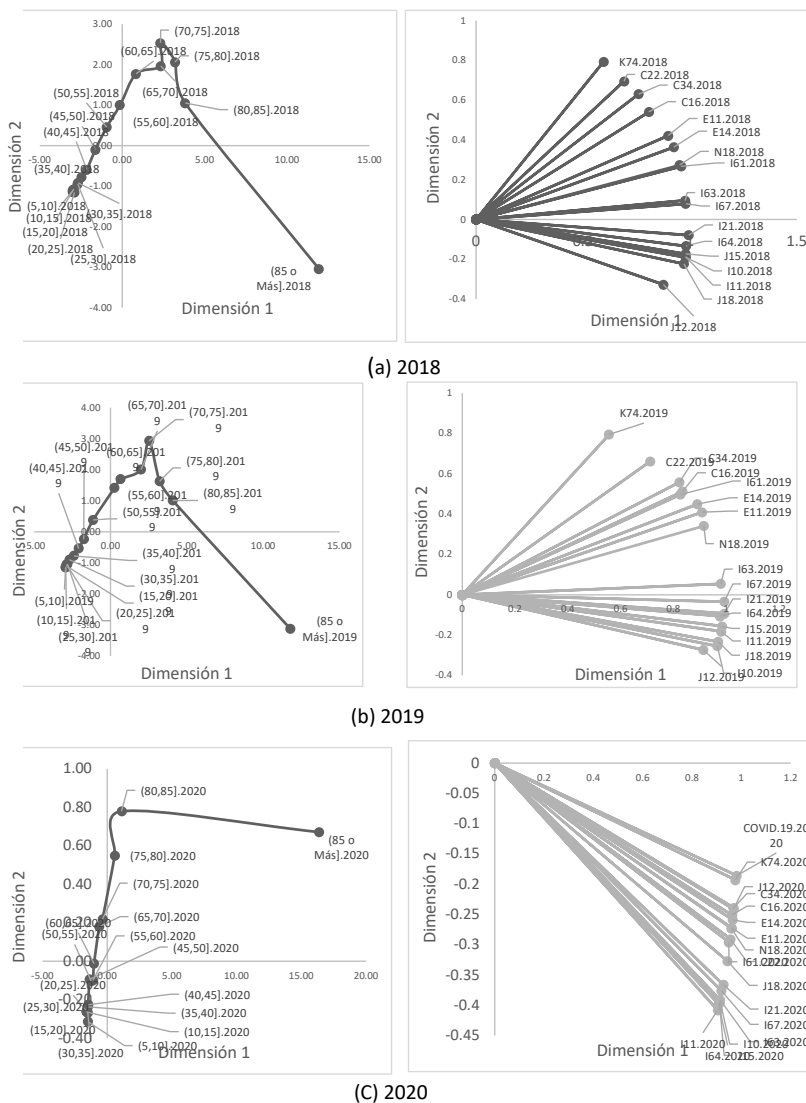
Right Plot Data (Approximate Coordinates):

Label	Dimensión 1 (X)	Dimensión 2 (Y)
K74.2020	0.9	-0.15
COVID.19.20	0.8	-0.25
J12.2020	0.7	-0.3
J1.2020	0.6	-0.35
C46.2020	0.5	-0.4
E11.2020	0.4	-0.45
E1.2020	0.3	-0.5
H18.2020	0.2	-0.55
M16.2020	0.1	-0.6
C34.2020	0.0	-0.65
I21.2020	-0.1	-0.7
I67.2020	-0.2	-0.75
I11.2020	-0.3	-0.8
I10.2020	-0.4	-0.85
I26.2020	-0.5	-0.9
I64.2020	-0.6	-0.95
I15.2020	-0.7	-1.0

Tumores malignos (C16, C22, C34)	Diabetes mellitus (E11, E14)	Hipertensión (I10,I11)	Influenza, gripe y neumonía (J12, J15, J18)	Covid-19
Fibrosis y cirrosis hepática (K74)	Insuficiencia Renal (N18)	Infarto Agudo de miocardio (I21)	Enfermedad cerebrovascular (I61, I63, I64, I67)	

70

Figura 7: Análisis de intraestructura en mujeres



Tumores malignos (C16, C22, C34)	Diabetes mellitus (E11, E14)	Hipertensión (I10, I11)	Influenza, gripe y neumonía (J12, J15, J18)	Covid-19
Fibrosis y cirrosis hepática (K74)	Insuficiencia Renal (N18)	Infarto Agudo de miocardio (I21)	Enfermedad cerebrovascular (I61, I63, I64, I67)	

Fuente: elaboración del autor usando datos de Defunciones Generales, 2018-2020, INEC.

Los resultados obtenidos a través del PTA, fueron validados con el PERMANOVA, el cual mostró que existen diferencias significativas entre los niveles de los efectos principales: año, edad, género; así también entre las interacciones de segundo orden: edad-año y edad-género ($\text{val-p} < 0.05$); no así con la interacción año-género ($F = 0.944$; $\text{val-p} = 0.443$), ni la interacción de tercer orden edad-año-género ($F = 0.303$; $\text{valor-p} = 1.000$) (Tabla 4).

Tabla 4: Resultados Permanova

Variación	Grados de libertad	Sumas cuadráticas	Medias Cuadráticas	F	Valor p	Sign.
Año	1	1.9098	1.90984	41.583	0.001	***
Edad	16	20.2468	1.26543	27.552	0.001	***
Gen	1	0.2729	0.27287	5.941	0.002	**
Año:Edad	16	1.6361	0.10226	2.226	0.001	***
Año:Gen	1	0.0433	0.04334	0.944	0.443	
Edad:Gen	16	1.4145	0.08841	1.925	0.003	**
Año:Edad:Gen	16	0.2223	0.0139	0.303	1	
Residuals	34	1.5616	0.04593			
Total	101	27.3075				

Fuente: elaboración del autor a partir de los datos de INEC.

Construcción de Indicadores a través de PTA

PTA utiliza descomposición de valores singulares de la matriz compromiso para poder construir los scores tanto para filas (edades) y columnas (causas de muerte). Estos scores para las filas y columnas tanto para hombres y mujeres se presentan en la Tabla 5. Se trabajó con solo una dimensión debido a que explicaba alrededor de 87 por ciento tanto para hombres como para mujeres, es decir un valor singular de 0.93.

Los índices se construyeron a través del producto matricial $\text{Score}_{\text{grupo etareo}} \text{Score}_{\text{causas}}^t$, es decir, se multiplicó *score* (proyección) en la dimensión de mayor variabilidad de los grupos etarios por la transpuesta del score en la dimensión de mayor variabilidad de las causas. De esa manera se obtienen unos indicadores que puntúan más alto mientras más edad tengas, además da mayores puntuaciones a las enfermedades cerebrovasculares, hipertensivas e infarto agudo de miocardio e insuficiencia renal crónica y menor peso a la Covid-19 porque estos indicadores se han

obtenido de la matriz compromiso que se obtiene combinando linealmente las matrices de los años 2018, 2019 y 2020.

Tabla 5: Primera Dimensión, Scores Hombres y mujeres

Grupo etario	Score Hombres	Score Mujeres	Causas	Score Hombres	Score Mujeres
(5,10]	-0,81	-0,66	C16	0,23	0,23
(10,15]	-0,82	-0,67	C22	0,24	0,21
(15,20]	-0,8	-0,66	C34	0,24	0,22
(20,25]	-0,78	-0,66	Covid-19	0,08	0,08
(25,30]	-0,76	-0,63	E11	0,23	0,24
(30,35]	-0,69	-0,61	E14	0,23	0,24
(35,40]	-0,64	-0,56	I10	0,26	0,25
(40,45]	-0,58	-0,51	I11	0,26	0,25
(45,50]	-0,44	-0,42	I21	0,26	0,26
(50,55]	-0,16	-0,29	I61	0,24	0,24
(55,60]	-0,01	-0,08	I63	0,26	0,25
(60,65]	0,29	0,05	I64	0,26	0,25
(65,70]	0,46	0,34	I67	0,26	0,26
(70,75]	0,72	0,41	J12	0,17	0,24
(75,80]	0,94	0,63	J15	0,25	0,25
(80,85]	0,94	0,81	J18	0,26	0,25
(85 o Más]	3,13	3,51	K74	0,19	0,19
			N18	0,26	0,25

Fuente: elaboración del autor a partir de los datos de INEC.

Estos indicadores se presentan en la Tabla 6 para hombres y mujeres. Para las categorías joven, mediana y mayor edad se han colocado los cuantiles de orden 20, 50 y 80 que corresponden a las edades menores a 45 años, entre 45 y 55 y mayores a 60.

Se puede apreciar que de manera general los indicadores son mayores en hombres que en mujeres. Siendo los indicadores similares para la categoría de jóvenes en ambos sexos, y el de mayor valor el de la Covid-19, que fue la causa que ocasionó la mayor cantidad de muertes para ese grupo de edad en relación a las causas analizadas en este estudio. Situación similar ocurre en la mediana edad, mientras que para la categoría “personas mayores”, aunque la Covid-19 fue la que ocasionó mayor cantidad de muertes en comparación con las demás causas analizadas y en todos los años

analizados, el infarto agudo de miocardio generó una cantidad similar de muertes en el 2020 y ha sido la que mayor cantidad de registros de muerte ha ocasionado en los años analizados, por eso tiene la mayor puntuación.

Tabla 6: Indicadores multivariantes Hombres y mujeres

Causas	Jóvenes	Hombres		Jóvenes	Mujeres	
		Mediana edad	Personas mayores		Mediana edad	Personas mayores
Tumores Malignos neoplasias	-0,064	-0,008	0,099	-0,061	-0,013	0,036
Diabetes mellitus	-0,067	-0,009	0,098	-0,068	-0,01	0,107
Enfermedades hipertensivas	-0,07	-0,007	0,111	-0,069	-0,011	0,108
Influenza, gripe y neumonía	-0,063	-0,007	0,099	-0,071	-0,012	0,107
Fibrosis y cirrosis del hígado	-0,052	-0,006	0,078	-0,051	-0,008	0,082
Insuficiencia renal crónica	-0,072	-0,006	0,112	-0,068	-0,011	0,109
Covid-19	-0,022	-0,002	0,034	-0,022	-0,004	0,036
Infarto agudo del miocardio	-0,073	-0,007	0,114	-0,072	-0,011	0,116
Enfermedades cerebrovasculares	-0,069	-0,008	0,109	-0,069	-0,01	0,115

Fuente: elaboración del autor a partir de los datos de INEC.

Como se mencionó antes, estos indicadores fueron obtenidos con la matriz compromiso, que ha dado mayor peso a las matrices de datos de los años 2018 y 2019, sería importante compararlos con los indicadores obtenidos solamente para 2020, por ello a continuación en la Tabla 7 se los presenta.

Los indicadores obtenidos para el año 2020 (Tabla 7) son diferentes y se alejan a los indicadores del compromiso, esto se debe al cambio de la intraestructura de 2020 en relación a los otros años (ver figuras 6 y 7), lo que refleja la situación fuera de control que se vivió por la pandemia de la Covid-19.

Tabla 7: Indicadores multivariantes Hombres y mujeres, año 2020

Causas	Hombres			Mujeres		
	Jóvenes	Mediana edad	Personas mayores	Jóvenes	Mediana edad	Personas mayores
Tumores Malignos neoplasias	-0,9933	-0,0520	0,9467	-1,0067	-0,1067	0,9367
Diabetes mellitus	-1,0300	-0,0610	0,9845	-1,0350	-0,0950	0,9940
Enfermedades hipertensivas	-0,9800	-0,0400	0,9600	-1,0060	-0,0985	0,9620
Influenza, gripe y neumonía	-0,9767	-0,0573	0,9433	-1,0133	-0,1077	0,9543
Fibrosis y cirrosis del hígado	-0,8000	-0,0600	0,7500	-0,8300	-0,0800	0,7900
Insuficiencia renal crónica	-1,0600	-0,0800	1,0070	-1,0600	-0,1000	1,0010
Covid-19	-0,9600	-0,0800	0,9000	-0,9180	-0,0890	0,8855
Infarto agudo del miocardio	-1,0400	-0,0700	0,9900	-1,0600	-0,1230	0,9740
Enfermedades cerebrovasculares	-1,0150	-0,0600	0,9850	-1,0530	-0,1153	0,9840

Fuente: elaboración del autor a partir de los datos de INEC.

DISCUSIÓN Y CONCLUSIONES FINALES

En este trabajo se aplica la metodología XSTATIS a datos de mortalidad de las principales causas de muerte de Ecuador durante los años 2018, 2019 y 2020. Se puede apreciar la gran utilidad de esta técnica multivariante que va desde analizar la interestructura, intraestructura de cada tabla de datos, hasta la construcción de indicadores que pueden resumir las características importantes para cada subpoblación de estudio.

En la interestructura se refleja el comportamiento atípico del año 2020 en relación a los anteriores, mientras que a través de los scores obtenidos de la matriz compromiso se refleja las relaciones entre causas principales de muerte y grupo de edad durante el período 2018-2020 (figuras 4 y 5) de manera resumida usando una menor dimensión de trabajo (dos dimensiones). Lo visualizado en la matriz compromiso puede ser contrastado con los conteos de muerte observados durante el período 2018-2020 (Tabla 3).

En las figuras 4 y 5 se formaron grupos o estratos de edades tanto en hombres como en mujeres que se pueden visualizar por las burbujas in-

corporadas en los gráficos. En mujeres se contabilizaron siete y nueve por ciento de casos en los grupos de los más jóvenes y de la mediana edad, respectivamente con valores negativos en la primera dimensión, esto se debe a que el conteo de casos en esos datos es menor. Se puede apreciar como a través de esta metodología se resaltan las similitudes entre filas y variables a través del estudio de la matriz compromiso y de la interestructura.

La matriz compromiso nos permitió observar las asociaciones que se mantuvieron estables a través del tiempo en cuanto a causas de muerte según los rangos de edad; indicando que las principales causas de muerte fueron Covid-19, infartos (I), influenza, gripe y neumonía (VR); y que en las edades de 45 a 55 años en hombres y de 50 a 60 años en mujeres las causas principales fueron tumores malignos (T) y diabetes mellitus (D).

La construcción de indicadores a través de variables latentes de la forma que se presenta en este trabajo permite ver la interacción causa y grupo a través de los años. Como es de esperarse la mayor cantidad de muertes se registran en edades mayores, sin embargo, en el año 2020 se incrementó las muertes en los grupos jóvenes y de mediana edad por lo que en los indicadores se reflejó que la Covid-19 tenía la mayor puntuación en estos grupos de edades para hombres y mujeres (Tabla 4), mientras que, en las edades mayores, el infarto agudo de miocardio fue la que obtuvo mayor puntuación. Estos indicadores al compararse con los obtenidos en el año 2020 reflejan que ese año fue atípico, por lo que los indicadores se pueden también usar para diseño de gráficos de control de calidad de variables que sea de interés monitorear para los investigadores.

Una vez más, a través de una aplicación a datos de mortalidad, se puede apreciar como la metodología XSTATIS es muy útil en diversos ámbitos de estudios, como en el presente trabajo, en el que se analizó la interacción de variables en diversos grupos a través del tiempo. Adicionalmente, para realizar el presente estudio no se requirió del total de subpoblaciones que, en el caso de Ecuador, se basan en proyecciones porque los censos se hacen cada diez años. Estas proyecciones se hacen de manera general y no se estiman para todos los posibles subgrupos de la población ecuatoriana (INEC, 2023). Métodos como regresión logística y Poisson requieren de los totales éxitos, fracasos, que para este caso (Población viva, Población muerta).

REFERENCIAS BIBLIOGRÁFICAS

Abdi, H., Williams, L. J., Valentin, D., & Bennani-Dosse, M., 2014, "STATIS and DISTATIS: Optimum multitable principal component analysis and three way met-

ric multidimensional scaling”, in *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2), 124–167. doi:10.1002/wics.19

Anderson, MJ., 2001, “A new method for non-parametric multivariate analysis of variance”, in *Austral Ecology*. 26 (1): 32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x.

Anderson, MJ, 2017, *Permutational Multivariate Analysis of Variance* (PERMANOVA). DOI:10.1002/9781118445112.stat07841. In book: Wiley StatsRef: Statistics Reference Online (pp.1-15). Massey University

Carrión-Bósquez, N., Castelo-Rivas, W., Alcívar-Muñoz, M., Quiñonez-Cedeño, L., & Llambo-Jami, H., 2022, “Influence of Covid-19 on the work environment of health workers in Ecuador”, en *Revista Información Científica*, 101(1), e3632. Epub 01 de enero de 2022. Recuperado en 04 de enero de 2024, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1028-99332022000100009&lng=es&tlng=en.

Castro, J. & Galindo, P., 2000, *Estadística Multivariante, Análisis de correlaciones*, Amaru Ediciones. España

Cevallos-Valdiviezo, H., Vergara-Montesdeoca, A., and Zambrano-Zambrano, G., 2021, “Measuring the impact of the covid-19 outbreak in ecuador using preliminary estimates of excess mortality, march 17–october 22, 2020”, in *International Journal of Infectious Diseases*, 104:297–299.

Cuéllar, L., Torres, I., Romero-Severson, E., Mahesh, R., Ortega, N., Pungitore, S., Ke, R., and Hengartner, N., 2021, “Assessing the impact of human mobility to predict regional excess death in Ecuador”, in *Scientific reports*, 12, 370.

Chen, F.-M., Feng, M.-C., Chen, T.-C., Hsieh, M.-H., Kuo, S.-H., Chang, H.-L., Yang, C.-J., and Chen, Y.-H., 2021, “Big data integration and analytics to prevent a potential hospital outbreak of Covid-19 in Taiwan”, in *Journal of Microbiology, Immunology and Infection*, 54(1):129–130

Fernández-Naranjo, R. P., Vásconez-González, E., Simbaña-Rivera, K., Gómez-Barreno, L., Izquierdo-Condoy, J. S., Cevallos-Robalino, D., and Ortiz-Pardo, E., 2021, “Statistical data driven approach of Covid-19 in ecuador: R0 and rt estimation via new method”, in *Infectious Disease Modelling*, 6:232–24

García-Bustos, S., Ruiz-Barzola, O., Cáceres, D., Márquez, KK, Valencia, L., Vergara, C., & González-Narváez, MA, 2021, “Análisis estadístico de las enfermedades asociadas a la mortalidad por Covid-19 en un Hospital de Ecuador durante el año 2020”, en *Bionatura*.

González-Narváez, M, 2021, *Contribuciones al análisis multivariante de matrices de datos multivía*, Universidad de Salamanca, España. DOI 10.14201/gre-dos.149381.

González-Narváez, M, Ruiz-Barzola, O. y Nieto-Librero, A, 2020, “Análisis multivariante: Un recorrido por las técnicas de reducción de dimensiones” en *Revista Matemática*, 18(2), disponible en <http://www.revistas.espol.edu.ec/index.php/matematica/article/view/747/791>

INEC, 2021, *Boletín Técnico, Registro estadístico de defunciones generales 2020*, Ecuador, disponible en https://www.ecuadorencifras.gob.ec/documentos/web-inec/Poblacion_y_Demografia/Defunciones_Generales_2021/Bolet%C3%ADn_EDG_v1.pdf.

INEC, 2022, *Registro estadístico de defunciones generales 2020*, Ecuador, disponible en <https://www.ecuadorencifras.gob.ec/defunciones-generales/>

INEC, 2023, *Proyecciones Poblacionales*, Ecuador Proyecciones Poblacionales | ([ecuadorencifras.gob.ec](https://www.ecuadorencifras.gob.ec))

Jiménez, JJ, Decaens, T & Rossi, JP, 2006, “Stability of the spatio-temporal distribution and niche overlap in neotropical earthworm assemblages”, in *Acta Oecologica*, 30:299–311.

Kroonenberg, P., 1989, “The analysis of multiple tables in factorial ecology. III: three-mode principle component analyse: Analyse triadique complète”, in *Acta Oecologica/Oecologica Generalis*, 10:245–256

Li, M., Dong, Y., Wang, H., Guo, W., Zhou, H., Zhang, Z., Tian, C., Du, K., Zhu, R., Wang, L., Zhao, L., Fan, H., Luo, S., & Hu, D., 2020, “Cardiovascular disease potentially contributes to the progression and poor prognosis of covid-19”, in *Nutrition, Metabolism and Cardiovascular Diseases*, 30:7.

Medina, J., 2015, *Análisis Multivariante del Uso del Tiempo de la Población Colombiana mediante el HJ Biplot y el Análisis Triádico Parcial*, Universidad de Salamanca, España.

Mera-Intriago E., 2023, *Salud y discapacidad en Ecuador: Factores Sociodemográficos en el contexto de la Transición Sanitaria*. Tesis doctoral no publicada. Universitat Autònoma de Barcelona.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H., 2022, “PERMANOVA-S: association test for microbial community composition that accommodates confounders and multiple distances” in *Bioinformatics*, 32(17), 2618-26251

Puchaicela, S., Loza, J., Fiallo, I., Benítez, A., Amaya, A., 2022, “Evaluación de estrés, ansiedad y depresión en Ecuador durante la pandemia de Covid-19” en *La Ciencia al Servicio de la Salud y la Nutrición*, 13(1):13-25. Disponible en <http://revistas.esPOCH.edu.ec/index.php/cssn/article/view/667>. Fecha de consulta el 04/01/2024, doi: <http://dx.doi.org/10.47244/cssn.Vol13.Iss1.667>.

R Core Team, 2022, *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Disponible en <https://www.R-project.org/>

Rodríguez, C., Galindo, P., Vicente, P., 2020, *Contribuciones a los métodos STATIS basados en técnicas de aprendizaje no supervisado*, Tesis doctoral no publicada. Universidad de Salamanca.

- Rolland, A., Bertrand, F., Maumy, M. & Jacquet, S., 2009, "Assessing phytoplankton structure and spatio-temporal dynamics in a freshwater ecosystem using a powerful multiway statistical analysis", in *Water Research*, 43:3155–3168
- Rossi, J.P., Nardin, M., Godefroid, M., Ruiz-Diaz, M., Sergent, A.S., Martinez-Meier, A., Pâques, L., Rozenberg, P., 2014, "Dissecting the space-time structure of tree-ring datasets using the partial triadic analysis", in *PLoS One*, 9:1–13
- Ruiz Cantero, M. T., 2021, "Las estadísticas sanitarias y la invisibilidad por sexo y de género durante la epidemia de Covid-19", in *Gaceta Sanitaria*, 35(1):95–98.
- Rygel, L., O'sullivan, D. & Yarnal, B., 2006, "A Method for Constructing a Social Vulnerability Index: An Application to Hurricane Storm Surges in a Developed Country", in *Mitigation and Adaptation Strategies for Global Change*, 11, 741–764 <https://doi.org/10.1007/s11027-006-0265-6>
- Slimani, N, Jiménez, JJ, Guilbert, E, Boumaïza, M, Thioulouse, J., 2020, "Surface water quality assessment in a semiarid Mediterranean region (Medjerda, Northern Tunisia) using partial triadic análisis", in *Environmental Science and Pollution Research International*, 27(24):30190-30198. doi: 10.1007/s11356-020-09326-7. Epub 2020 May 25. PMID: 32451888.
- Thioulouse, J., 2011, "Simultaneous analysis of a sequence of paired ecological tables: A comparison of several methods", in *The Annals of Applied Statistics*, 5, 2300–2325
- Thioulouse J, Chessel D, 1987, "Les analyses multitableaux en écologie factorielle. I: De la typologie d'état à la typologie de fonctionnement par l'analyse triadique", in *Acta Oecologica*, 8:463–480
- Thioulouse, J, Simier, M, Chessel, D, 2004, "Simultaneous analysis of a sequence of paired ecological tables", in *Ecology*, 85:272–283
- Thioulouse, J, Dray, S, Dufour, AB, Siberchicot, A, Jombart, T, Pavoine, S, 2018, *Multivariate analysis of ecological data with ade4*. 330p, Springer, New York
- Vera, F., Solórzano, M., Ochoa, G., García-Bustos, S. and Cevallos, S., 2019, "Tablas de Mortalidad de Ecuador continental mediante un análisis de supervivencia", in *Papeles de Población*, vol. 24 No. 97. Disponible en <https://rppoblacion.uaemex.mx/article/view/9061>. Fecha de consulta el 20/03/2019.
- Vyas, S. and Kumaranayake, L., 2006, "Constructing socio-economic status indices: how to use principal components análisis", in *Health Policy and Planning*, 21(6):459–468.
- WHO, 2022, *Coronavirus disease (Covid-19) situation reports jun/2022*.
- WHO, 2019, *International statistical classification of diseases and related health problems (icd)*, Available at <https://www.who.int/classifications/classification-of-diseases>
- Wu, Z, & McGoogan, JM, 2020, "Características y lecciones importantes del brote de la enfermedad por coronavirus 2019 (Covid-19) en China: resumen de un infor-

me de 72 314 casos del Centro chino para el control y la prevención de enfermedades”, in *JAMA*, Publicado en línea el 24/01/2020. doi: 10.1001 / jama.2020.2648

Zerna-Bravo, C., Alvarado-Álvarez, A., Herrera-Rivera, J., & Martin, F., 2022, “Identificación de riesgos obstétricos en tiempos de Covid-19 en un área comunitaria de salud, Ecuador”, en *Revista Información Científica*, 101(3), e3847. Epub 09 de mayo de 2022. Consultado el 04/01/2024, disponible en http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1028-99332022000300010&lng=es&tln-g=es.

Zhang, J., Wang, X., Jia, X., Li, J., Hu, K., Chen, G., Wei, J., Gong, Z., Zhou, C., Yu, H., Yu, M., Lei, H., Cheng, F., Zhang, B., Xu, Y., Wang, G., & Dong, W., 2020, “Risk factors for disease severity, unimprovement, and mortality in Covid-19 patients in Wuhan, China”, in *Clinical Microbiology and Infection*, 26(6):767–772

RESUMEN CURRICULAR DE LOS AUTORES

Sandra García Bustos

Es profesora de Estadística en la Universidad ESPOL (Escuela Superior Politécnica del Litoral) en Guayaquil, Ecuador. En 2014, recibió su Doctorado en Estadística y Optimización en la Universidad Politécnica de Valencia, España. Tiene estudios de tercer nivel en Estadística y Maestría en Administración de Empresas. Sus intereses de investigación actuales son el control estadístico de procesos y la Estadística Social.

Dirección electrónica: slgarcia@espol.edu.ec

Registro ORCID: <https://orcid.org/0000-0003-2978-0330>

Omar Ruiz Barzola

Doctor en Estadística y Optimización. Profesor Investigador desde 2005 de la ESPOL; y desde 2017 profesor asociado del Departamento de Estadística de la Universidad de Salamanca, España. Ex-director del Centro de Estadística de la ESPOL. Presidente del International Biometric Society (IBS) capítulo Ecuador. Experto en la aplicación de Estadística en las ciencias de la vida y la salud, autor o coautor de aproximadamente 50 aportaciones científicas; y director o investigador de más de una docena de proyectos de investigación.

Dirección electrónica: oruiz@espol.edu.ec

Registro ORCID: <https://orcid.org/0000-0001-8206-1744>

Miguel Morán

Es Ingeniero en Estadística de la Universidad ESPOL. Trabaja en el Departamento de Matemáticas de la ESPOL. Sus intereses de investigación actuales son el control estadístico de procesos y la Estadística Social.

Dirección electrónica: mianmora@espol.edu.ec

William Tandazo

Es Ingeniero en Estadística de la Universidad ESPOL. Trabaja en el Departamento de Matemáticas de la ESPOL. Su intereses de investigación actual es la Estadística Social.

Dirección electrónica: wtandazo@espol.edu.ec

Artículo recibido el 01 de noviembre de 2022 y aceptado el 24 de noviembre de 2023



Disponible en:

<https://www.redalyc.org/articulo.oa?id=11281623004>

Cómo citar el artículo

Número completo

Más información del artículo

Página de la revista en redalyc.org

Sistema de Información Científica Redalyc
Red de revistas científicas de Acceso Abierto diamante
Infraestructura abierta no comercial propiedad de la
academia

Sandra García-Bustos, Omar Ruiz-Barzola, Miguel Morán,
William Tandazo

**Método XSTATIS para estudio de principales causas de
muerte en Ecuador 2020**

**XSTATIS method for the study of the main causes of death
in Ecuador 2020**

Papeles de población

vol. 29, núm. 116, p. 53 - 81, 2023

Universidad Autónoma del Estado de México, Centro de
Investigación y Estudios Avanzados de la Población,

ISSN: 1405-7425

ISSN-E: 2448-7147

DOI: <https://doi.org/10.22185/24487147.2023.116.13>