



Población y Salud en Mesoamérica

Revista electrónica publicada por el
Centro Centroamericano de Población,
Universidad de Costa Rica, 2060 San José, Costa Rica
<http://ccp.ucr.ac.cr>

Población y Salud en Mesoamérica

Revista electrónica semestral, ISSN-1659-0201

Volumen 10, número 2, artículo 4

Enero - junio, 2013

Publicado 1 de enero, 2013

<http://ccp.ucr.ac.cr/revista/>

Análisis factorial para el estudio de la mortalidad de Costa Rica. Periodo 1900-2010

Eduardo Aguilar Fernández



Protegido bajo licencia Creative Commons

Centro Centroamericano de Población

Análisis factorial para el estudio de la mortalidad de Costa Rica. Periodo 1900-2010

Factor analysis for the study of mortality in Costa Rica. Period 1900-2010

Eduardo Aguilar Fernández¹

RESUMEN

La mortalidad es uno de los componentes de la dinámica demográfica constituido en un importante indicador del reflejo del estado de salud de una población, por lo que conocer su comportamiento es de suma importancia en el desarrollo de las sociedades. El objetivo de la presente investigación es construir un modelo de análisis factorial mediante la aplicación de la técnica de componentes principales a un conjunto de tasas de mortalidad por grupo de edad y sexo para identificar componentes influyentes en la mortalidad de Costa Rica durante el período 1900-2010. El análisis factorial permitió identificar inicialmente tres factores principales que explican en un 96,6 % la variabilidad de la mortalidad de Costa Rica. Se espera que esta serie de estimaciones contribuyan al estudio del comportamiento de la mortalidad y de sus implicaciones para el desarrollo del país y, a su vez, sirva de base para que futuros estudios complementen los alcances de la presente investigación.

Palabras clave: tasa de mortalidad, esperanza de vida, análisis factorial, componentes principales.

ABSTRACT

The mortality is one of the components of the demographic dynamics constituted as an essential pointer of a population's health status reflection, that's why it is extremely important for the societies' development to know about its behavior. The objective of the current investigation is to build a factor analysis model by applying the technique of main components to a death rate set, according to age and gender groups, in order to identify influential components in Costa Rica's mortality between 1900 and 2010. The factor analysis allowed us to initially identify three main factors which show in a 96,6 % the variability in Costa Rica's mortality. It is expected from these estimates to contribute with the study of mortality, its behaviors and implications for the development of the country. At the same time, these estimates are expected to work as a base, with the purpose of complementing further studies with the achievements of this research.

Key words: death rate, life expectancy, factor analysis, main components.

Recibido: 28 set. 2012

Aprobado: 5 oct. 2012

¹ Profesor, Universidad Nacional y Universidad de Costa Rica. COSTA RICA. eaguilar2@gmail.com

1. INTRODUCCIÓN

La mortalidad refleja el estado de salud de una población y, además, es uno de los componentes de la dinámica demográfica. Los cambios en su comportamiento modifican el patrón demográfico de una población (Bongaarts, 2005). Por ello, su estudio y conocimiento adquiere gran importancia, ya que aspectos como el nivel de mortalidad, la estructura por edad y sexo, así como la composición por causa de muerte se utilizan como indicadores de salud y bienestar (Peláez, 1998).

Otro aspecto fundamental es que su comportamiento puede presentar diferencias importantes entre países, regiones, clases sociales; incluso, dentro de una misma zona puede variar conforme la edad o el sexo. Conocer estas diferencias es un primer paso fundamental para entender su comportamiento y las causas que la determinan. Según Naciones Unidas (1962) uno de los objetivos de los estudios que analizan la mortalidad es poner de manifiesto los factores que determinan su nivel.

En este sentido, Ledermann y Breas (1959) utilizan el análisis factorial para el estudio de la mortalidad. Este análisis fue aplicado a las tasas de mortalidad por edad y sexo de diferentes países y en diversas épocas. Los datos concernientes al estudio provienen de 157 tablas de mortalidad correspondientes a unos 50 países y que están distribuidas a lo largo de la primera mitad del siglo XX.

De cada tabla se obtuvo 38 coeficientes representados por el logaritmo de los cocientes de mortalidad de hombres y mujeres de 18 grupos de edades (0-1, 1-4, 5-9, ..., 80-84) y la esperanza de vida al nacer (e_0) mediante la expresión $100 - e_0$. En total, se tienen 5 966 coeficientes variantes en el tiempo y el espacio e influenciados por diversos factores de manera que con ayuda del análisis puedan reconstruirse las variaciones y covariaciones de dichos coeficientes por medio de un pequeño número de variables subyacentes. Según lo menciona Naciones Unidas (1962), el análisis por sí solo no permite identificar las variables, pero al mostrar cómo se comportan en sus efectos permite orientar la investigación para su posible identificación.

El desarrollo de la investigación permitió identificar, inicialmente, tres componentes principales independientes y que pueden estudiarse por separado. El porcentaje de varianza total explicado por el modelo de tres factores es de 92,6%.

El primer índice o factor identificado por Ledermann y Breas (1959) explica el 77% de la dispersión de los logaritmos de los cocientes de mortalidad y su mayor influencia se da sobre la esperanza de vida y la mortalidad de las edades de 5 a 35 años. Los autores expresan que este factor refleja la mejora de las condiciones sanitarias generales.

El segundo factor explica solamente el 10% de la dispersión de los cocientes de las edades; sin embargo, para las edades superiores a los 40 años explica un 26% de la dispersión de los cocientes de los hombres y 13% para los de las mujeres. Por esta razón, este índice está relacionado con una parte de la mortalidad de los adultos.

El tercer factor, considerado independiente de los otros dos, explica solamente el 6,5% de la dispersión de los cocientes de las distintas edades. Su influencia se ve reflejada estrictamente en las últimas edades de la vida donde a su contribución es superior al 50% cuando la edad es de 70 años o más.

Finalmente, concluyen que la utilización de estos tres componentes principales es insuficiente para la estimación de los cocientes de mortalidad, por lo que recomiendan la incorporación de dos componentes adicionales, que expliquen las variaciones marginales locales, incluyendo la mortalidad infantil y las diferencias particulares en la mortalidad por sexo a partir de los 25 años.

Por otro lado, Naciones Unidas (1962) aplica, como complemento al estudio de Ledermann y Breas, el método de manera un poco distinta a datos provenientes de 26 países para períodos relativamente distintos.

De esta forma, el presente estudio pretende desarrollar un modelo de análisis factorial mediante la aplicación de la técnica de componentes principales para identificar componentes influyentes en la mortalidad de Costa Rica. Asimismo, se evalúa el desempeño del modelo mediante la comparación de las tasas de mortalidad estimadas con los valores observados en el período de estudio.

2. DATOS Y MÉTODOS

Para la presente investigación, la población de estudio está conformada por 4 218 tasas de mortalidad de Costa Rica definidas por sexo, edad y año, que comprenden el período que va desde el año 1900 hasta el año 2010. Como unidad de estudio se define la tasa de mortalidad, por sexo, del grupo de edad x en el año t .

Los datos del estudio corresponden a tasas de mortalidad, definidas para cada sexo, correspondientes al período 1900-2010, determinadas en forma anual y según grupo de edad.

La base de datos se elabora con las tasas correspondientes al período 1900-1950 proveniente de estimaciones realizadas por el historiador Héctor Pérez (Pérez, 2010) y los correspondientes al período 1950-2010, provenientes de las tablas de mortalidad elaboradas por el Centro Centroamericano de Población (CCP) en conjunto con el Instituto Nacional de Estadística y Censos (INEC) (CCP/INEC, 2008).

Las variables consideradas en el estudio son la tasa de mortalidad, el sexo, el grupo de edad y el año, así como el logaritmo natural de la tasa de mortalidad. A continuación se describen brevemente cada de ellas.

Tasa de mortalidad: hace referencia a la tasa específica de mortalidad por grupo de edad. La tasa específica de mortalidad está dada por el cociente del número de defunciones de personas con edad entre x y $x + n$ ocurridas en el año entre el tiempo vivido por la población de dicho grupo de edad en el mismo año.

Año: corresponde al año en que se determina la tasa de mortalidad. En este caso se tiene información anual de las tasas de mortalidad correspondientes al período 1900-2010.

Sexo: la información de mortalidad obtenida está clasificada por sexo y para su análisis se seguirá dicha clasificación.

Grupo de edad: corresponde al grupo de edades para la cual se encuentra estimada la tasa de mortalidad. La información se encuentra definida para el grupo de los menores de un año (0), luego los de edad de 1 a 4 años y de ahí adelante, considerando grupos quinquenales. De esta forma, los grupos de edad definidos son 0, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85 y más años.

Técnica de estudio (análisis factorial de las tasas de mortalidad por edad y sexo).

Como se describió en una sección anterior, en 1959, Ledermann y Breas aplicaron el análisis factorial a un conjunto de datos provenientes de 157 tablas de mortalidad de unos 50 países y correspondientes a la primera mitad del siglo XX.

En el presente estudio, la aplicación de la metodología de Ledermann y Breas se lleva a cabo con un grupo de 40 variables, 20 para cada sexo, provenientes de 111 tablas de vida con información referente a la mortalidad de Costa Rica para el periodo 1900-2010. Las variables están definidas por el logaritmo natural de la expresión $100 - e_0$, donde e_0 representa la esperanza de vida a nacer y el logaritmo natural de la tasa de mortalidad de los grupos de edad 0-1, 1-4, 5-9 y restantes grupos quinquenales hasta el grupo de 85 y más años para cada uno de los sexos.

Las variables descritas se denotarán por $x_1, x_2, \dots, x_{20}, x_{21}, \dots, x_{40}$ donde x_1 representa el logaritmo natural de la expresión $100 - e_0$ para el sexo femenino, x_2 representa el logaritmo natural de la tasa de mortalidad infantil femenina, x_3 representa el logaritmo natural de la tasa de mortalidad del grupo 1-4 femenino y así sucesivamente hasta x_{20} , la cual representa el logaritmo natural de la tasa de mortalidad del grupo de 85 años y más. Seguidamente, x_{21} representa el logaritmo natural de la expresión $100 - e_0$ para el sexo masculino, x_{22} representa el logaritmo natural de la tasa de mortalidad infantil masculina hasta, x_{40} representa el logaritmo natural de la tasa de mortalidad del grupo 85 años y más masculino.

El modelo factorial

Según lo explica Hernández (1998), el análisis factorial tiene como objetivo la reconstrucción de las variaciones y covariaciones de un conjunto de variables a través de un grupo, relativamente pequeño, de variables subyacentes a las que se le denominan factores, que no son medibles u observables, directamente, de manera que su identificación facilite la descripción de los datos originales. De esta forma, la metodología supone que un conjunto de variables x_i pueden representarse de la forma

$$\begin{aligned}
 x_1 - \mu_1 &= a_{11}f_1 + a_{12}f_2 + \dots + a_{1q}f_q + \varepsilon_1 \\
 x_2 - \mu_2 &= a_{21}f_1 + a_{22}f_2 + \dots + a_{2q}f_q + \varepsilon_2 \\
 &\vdots \\
 x_p - \mu_p &= a_{p1}f_1 + a_{p2}f_2 + \dots + a_{pq}f_q + \varepsilon_p
 \end{aligned}
 \tag{1}$$

Las expresiones f_1, f_2, \dots, f_q representan un conjunto de q (q mucho menor que p) variables aleatorias no observables llamadas factores comunes y $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ son llamados factores específicos o errores que representa otras fuentes de variación propias de cada variable. Los valores constantes a_{ik} , $i=1, 2, \dots, p$ y $k=1, 2, \dots, q$, representan las cargas factoriales de la variable i sobre el factor k (Hernández, 1998).

El principio del análisis factorial expresa que si Σ representa la matriz de varianzas y covarianzas, esta puede escribirse como

$$\Sigma = L\Phi L' + \Psi
 \tag{2}$$

Y que además

$$\text{cov}(X, F) = L\Phi
 \tag{3}$$

Si se supone $\Phi = I$ se obtiene el modelo ortogonal de factores donde entonces es posible definir que

$$\text{cov}(x_i, f_k) = a_{ik}
 \tag{4}$$

Y

$$\text{var}(x_i) = a_{i1}^2 + a_{i2}^2 + \dots + a_{iq}^2 + \psi_i
 \tag{5}$$

En la expresión anterior, a la suma de los cuadrados de las cargas factoriales a_{ik} se le llama comunalidad para la variable i y suele denotarse por h_i^2 . Esta expresión indica el porcentaje de varianza total de la variable i que explica el modelo factorial. Cuando se está utilizando un modelo ortogonal de factores y las variables están estandarizadas se tiene que $\text{cov}(x_i, f_k) = \text{corr}(x_i, f_k) = a_{ik}$, por lo que la comunalidad de la variable i puede escribirse como

$$\begin{aligned}
 h_i^2 &= \text{corr}^2(x_i, f_1) + \text{corr}^2(x_i, f_2) + \dots + \text{corr}^2(x_i, f_q) \\
 &= a_{i1}^2 + a_{i2}^2 + \dots + a_{iq}^2
 \end{aligned}
 \tag{6}$$

La correlación al cuadrado entre la variable i y el factor k ($corr^2(x_i, f_k)$) indica el porcentaje de varianza total de la variable i que explica el k -ésimo factor o componente. Asimismo, para dos variables cualesquiera x_i y x_j , la correlación entre ellas puede expresarse por

$$\begin{aligned} corr(x_i, x_j) &= corr(x_i, f_1)corr(x_j, f_1) + \dots + corr(x_i, f_q)corr(x_j, f_q) \\ &= a_{i1}a_{j1} + a_{i2}a_{j2} + \dots + a_{iq}a_{jq} \end{aligned} \quad (7)$$

Para determinar las cargas factoriales del modelo se utilizará el método de componentes principales. Bajo este procedimiento, cuando se usa la matriz de varianzas y covarianzas S o la matriz de correlaciones R , para un número q de factores comunes, las cargas factoriales que componen la matriz L se estiman por

$$\hat{L}_q = (\hat{a}_{ik}) = \left(\sqrt{\hat{\lambda}_1} \hat{e}_1, \dots, \sqrt{\hat{\lambda}_q} \hat{e}_q \right) \quad (8)$$

Donde $(\hat{\lambda}_1, \hat{e}_1), \dots, (\hat{\lambda}_q, \hat{e}_q)$ con $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_q$ representan los pares de raíces y vectores característicos asociados de la matriz S o R . De esta manera, el porcentaje de varianza total de la variable i (comunalidad), explicado por el modelo factorial se estima por

$$\hat{h}_i^2 = \hat{a}_{i1}^2 + \hat{a}_{i2}^2 + \dots + \hat{a}_{iq}^2 \quad (9)$$

Por su parte, la estimación de la varianza específica de la variable i se obtiene por

$$\hat{\psi}_i = \text{var}(x_i) - \sum_{k=1}^q \hat{a}_{ik}^2 \quad (10)$$

En cuanto a la selección del número de factores, dado que se utiliza el método componentes principales para su estimación y la factorización se lleva a cabo sobre la matriz de correlación R , se toma q igual al número de raíces características de R que son mayores a 1.

Rotación de la solución factorial

Hernández (1998) menciona que si se cuenta con una solución factorial inicial esta no es única, por lo que es posible hallar otra, multiplicando la primera solución por una matriz ortogonal, obteniendo así una transformación ortogonal que es conocida como rotación factorial. Esta rotación hace que las cargas factoriales del modelo inicial se transformen en nuevas cargas que reproducen en forma equivalente la matriz de correlaciones o de covarianzas y que además permiten una mejor interpretación de los factores o componentes que se obtienen. La rotación no modifica la bondad del ajuste inicial del modelo, por lo que la comunalidad y el porcentaje de varianza explicado por todos los factores no cambian pero si lo hace el porcentaje de varianza explicado en forma individual por cada uno de los factores.

Para llevar a cabo la rotación factorial, se utiliza el método varimax, el cual, según lo expresa Hernández (1998), tiene el fin de facilitar la interpretación de los factores por medio de la minimización del número de variables que presentan cargas altas en un determinado factor.

Bondad del ajuste

Para medir la pertinencia del análisis factorial se utilizarán varias pruebas estadísticas. La primera de ellas es la medida de adecuación muestral KMO (Kaiser-Meyer-Olkin). Esta medida permite comparar las correlaciones parciales entre las variables para ver si son suficientemente pequeñas (Hernández, 1998; Pardo & Ruiz, 2002). Su valor oscila entre 0 y 1 y según Kaiser (1974), citado por Hernández (1998), valores de 0,8 se consideran meritorios y de 0,9 se consideran maravillosos para justificar la realización de un análisis factorial. Para su cálculo se utiliza la expresión

$$KMO = \frac{\sum \sum r_{ij}^2}{\sum \sum r_{ij}^2 + \sum \sum s_{ij}^2}, \quad i \neq j \quad (11)$$

En la expresión anterior, r_{ij} representa el coeficiente de correlación simple entre las variables i y j . Por su parte, s_{ij} representa la correlación parcial entre la variable i y j eliminado el efecto de las restantes variables incluidas en el análisis.

También, se utilizará la prueba de esfericidad de Bartlett, la cual contrasta la hipótesis nula de que la matriz de correlaciones es una matriz identidad, esto es, que no existen correlaciones significativas entre las variables y, en consecuencia, el análisis factorial no es adecuado. El estadístico de Bartlett se distribuye, aproximadamente, con distribución de probabilidad chi-cuadrado (Pardo & Ruiz, 2002). Hernández (1998) apunta que valores altos del estadístico y una significancia asociada pequeña indican que la matriz de correlaciones es diferente de la identidad lo cual justifica el análisis factorial.

Finalmente, se estimará los residuos o diferencias entre la correlación estimada y la correlación observada. Según Hernández (1998), la magnitud de estos residuos permite evaluar la calidad de ajuste del modelo, ya que una elevada proporción de valores residuales grandes no contribuye a la validez del modelo ajustado.

Estimación de las tasas de mortalidad, la esperanza de vida al nacer

Para llevar a cabo estimaciones de la esperanza de vida al nacer y de la tasa de mortalidad del grupo de edad x se hace uso de la rotación factorial. Las nuevas cargas factoriales obtenidas a partir de esta solución permiten estimar para un año j el valor estandarizado de la variable i mediante la ecuación

$$\hat{z}_{ij} = \hat{A}_{i1}F_{1j} + \hat{A}_{i2}F_{2j} + \dots + \hat{A}_{iq}F_{qj}, \quad 1 \leq i \leq 40 \quad (12)$$

Dado que $\hat{x}_{ij} = \hat{z}_{ij}\sigma_i + \mu_i$, la esperanza de vida al nacer y la tasa de mortalidad del grupo de edad x pueden estimarse por medio del antilogaritmo de la expresión

$$\hat{x}_{ij} = \left(\hat{A}_{i1}F_{1j} + \hat{A}_{i2}F_{2j} + \dots + \hat{A}_{iq}F_{qj} \right) \sigma_i + \mu_i \quad (13)$$

Medidas de precisión

Anteriormente, se mencionó que una rotación no afecta la bondad del ajuste del modelo factorial, ya que la comunalidad y el porcentaje de varianza total explicada por los factores no cambia. Si se considera la solución rotada, la precisión de la estimación de la variable estandarizada i la representaremos por medio de la ecuación

$$r_{z_i \hat{z}_i}^2 = \hat{A}_{i1}^2 + \hat{A}_{i2}^2 + \dots + \hat{A}_{iq}^2 \quad (14)$$

Como las variables estandarizadas tienen varianza 1, la varianza específica de la variable i corresponde a

$$s_i^2 = 1 - r_i^2 \quad (15)$$

De esta forma, Ledermann y Breas (1959) mencionan que el error medio en la estimación del logaritmo de la variable i puede definirse por

$$S_i = \sigma_i \cdot s_i \quad (16)$$

Finalmente, a partir del antilogaritmo de la expresión anterior, es posible obtener el error medio relativo de la variable i mediante la fórmula siguiente

$$e_i = \exp(S_i) - 1 \quad (17)$$

Para la estimación del modelo se utilizará el paquete estadístico Stata 10.0 (Stata Corp.)

3. RESULTADOS

A continuación, se presentan los valores descriptivos de las variables en estudio, el modelo factorial ajustado con sus respectivas medidas de bondad de ajuste, así como estimaciones de las tasas de mortalidad por grupo de edad y de la esperanza de vida al nacer para cada sexo, según los procedimientos descritos en la sección metodológica.

Valores descriptivos y correlaciones iniciales

La mortalidad de Costa Rica durante los 110 años de estudio presenta el comportamiento natural de este componente demográfico, es decir, valores altos en los primeros años de vida (0 a 4), luego desciende para las edades de 5 a 14, de manera que en las edades posteriores inicia un comportamiento ascendente, el cual se mantiene hasta alcanzar niveles elevados en las edades avanzadas (Cuadro 1).

Por otro lado, la esperanza de vida al nacer en cada uno de los sexos está altamente correlacionada con la mayor parte de los grupos de edad exceptuando los de 75 y más años. Además, para cada sexo, las tasas de mortalidad de los distintos grupos de edad muestran correlaciones altas entre sí, excepto en los de edades avanzadas, donde la fuerza disminuye y para un grupo de edad particular, se observan correlaciones elevadas con sus grupos vecinos (Cuadros 2 y 3). También es importante mencionar que al correlacionar las variables de un sexo con las del otro, puede observarse la existencia de una fuerte relación entre las variables del mismo grupo (Cuadro 4).

El modelo factorial y medidas de bondad de ajuste

Como se detalló en la sección de métodos, el número de factores por retener será igual a la cantidad de raíces características mayores que 1. Los resultados revelan la presencia de tres raíces características mayores que 1, por lo que se escoge un modelo de tres factores. Esta cantidad concuerda con la obtenida inicialmente por Ledermann y Breas (1959), quienes identificaron tres componentes principales para el periodo y países contemplados por las tablas de mortalidad que trataron (aunque ellos mismos plantean la necesidad de considerar una cuarta y posiblemente una quinta componente para agotar las variables subyacentes capaces de describir la mortalidad). De esta manera, las variaciones y covariaciones de la mortalidad de Costa Rica para el periodo 1900-2010 pueden explicarse por medio de tres factores.

En cuanto a la bondad de ajuste del modelo, la prueba KMO (descrita en la sección metodológica) determina un valor de 0,9571, el cual indica que el análisis factorial es adecuado.

La prueba de esfericidad de Bartlett revela una valor chi-cuadrado de 17 119 con una significancia asociada de $p = 0,0000$, por lo que se considera que el análisis factorial es adecuado.

Por otro lado, los valores residuales entre las correlaciones estimadas y las correlaciones observadas, considerando solo los tres primeros factores del modelo, son bastantes pequeños y sólo 2 (0,054 y -0,104) de 780 residuos (0,26%) presentan un valor absoluto mayor a 0,05, por lo que se considera que el modelo ajusta bien (Cuadro 5).

Coefficientes de correlación y representación algebraica del modelo

Los valores a_1 (Cuadro 6) muestran la correlación entre el factor 1 y cada una de las variables del estudio. Se observa que la correlación entre dicho factor 1 y el logaritmo de $100 - e_0$ es bastante elevada, 0,9972 en el sexo femenino y 0,9978 en el masculino. En general, este factor 1 está altamente correlacionado con casi todas las variables, por lo que puede considerarse un índice o factor general de mortalidad. Por su parte, el factor 2 presenta una correlación alta con el grupo de 80-84 (0,6058 en el sexo femenino y 0,7866 en el masculino). El factor 3 está más asociado con la mortalidad de los grupos de edades de 75-79 y 85 y más.

Los valores a_1 , a_2 y a_3 descritos anteriormente permiten escribir las variables estandarizadas como combinaciones lineales de las variables aleatorias no observables o factores. Por lo tanto, la solución del modelo factorial ortogonal para algunas de las variables puede escribirse como

Sexo femenino

$$\ln(100 - e_0): \hat{z}_{1j} = 0,9972f_{1j} - 0,0121f_{2j} + 0,0178f_{3j}$$

$$\ln(m(0,1)): \hat{z}_{2j} = 0,9805f_{1j} + 0,0526f_{2j} - 0,0043f_{3j}$$

Sexo Masculino

$$\ln(100 - e_0): \hat{z}_{21j} = 0,9978f_{1j} - 0,0382f_{2j} + 0,0233f_{3j}$$

$$\ln(m(0,1)): \hat{z}_{22j} = 0,9776f_{1j} + 0,0572f_{2j} - 0,0199f_{3j}$$

Porcentaje de varianza explicado

El porcentaje de varianza total explicado por los tres factores es de 96,62%. De este porcentaje, el 87,17% es explicado por el primer factor, el 6,03% y el 3,42% son explicados por el segundo y tercer factor, respectivamente (Cuadro 6).

Por otro lado, el modelo de tres factores explica el 99,49% de la varianza total de la variable $\ln(100 - e_0)$ femenina (cuadro 6) y el 99,44% ($0,9972^2$) de la misma es explicada por el factor 1. En cuanto al logaritmo natural de la mortalidad infantil, se tiene que el 96,42% de la varianza total es explicada por el modelo factorial (cuadro 6) y que el 96,13% de dicha varianza la explica el factor 1.

Para el sexo masculino, el modelo de tres factores explica el 99,76% (cuadro 6) de la varianza total de la variable $\ln(100 - e_0)$ y el $(0,9978)^2 = 99,56\%$ de esta varianza es explicada por el factor 1. Para el logaritmo natural de la mortalidad infantil, se tiene que el 95,94% (cuadro 6) de la varianza total es explicada por el modelo y que el 95,57% de dicha varianza la explica el factor 1.

También es destacable que para las edades entre los 20 y 54 años en el sexo femenino el porcentaje de varianza total explicado por el modelo supera el 99%, siendo el grupo de 30-34 donde el aporte es mayor (99,42%) (cuadro 6).

En términos generales, el porcentaje de varianza total de cada variable explicado por el modelo es mayor al encontrado por Ledermann y Breas (1959) en la mayoría de los grupos de edad. Por otro

lado, en el caso de la mortalidad infantil y en la del grupo de 1 a 4 años la contribución en este porcentaje no sobrepasa el 97%.

Rotación del modelo factorial

Como se mencionó en la sección metodológica, las cargas factoriales del modelo inicial pueden ser transformadas en nuevas cargas que reproducen en forma equivalente la matriz de correlaciones o de covarianzas y que, además, permiten una mejor interpretación de los factores o componentes que se obtienen. La rotación del modelo factorial genera nuevas cargas A_1 , A_2 y A_3 que indican que el primer factor está asociado fuertemente con los grupos de edad menores a los 75 en ambos sexos, además el factor 2 se asocia con mayor fuerza a la mortalidad del grupo de 75 a 79 en ambos sexos y la mortalidad masculina del grupo de 80 a 84, mientras que el tercer factor tiene mayor asociación con el grupos abierto de 85 y más años en ambos sexos y con la mortalidad femenina del grupo de 80 a 84 (Cuadro 6).

Las nuevas cargas factoriales permiten observar que el porcentaje de varianza explicado por el primer factor es de 77,42%, mientras que los factores 2 y 3 explican, respectivamente, el 14,34% y el 4,86% de la varianza total (Cuadro 6). Además, se tiene que para ambos sexos, el factor 2 explica alrededor de un 15% de la variabilidad de la mortalidad de 0 a 14 años, alrededor de un 2% en los grupos de 15 a 39, mientras que a partir de los 40 años este factor explica alrededor del 19% de la variabilidad. El factor 3 no aporta mucho en la explicación de la variabilidad de los logaritmos en las edades menores a 80 en ambos sexos, pues su contribución no alcanza el 1% en promedio.

Con la rotación, el porcentaje de varianza total de cada variable explicado individualmente por cada factor cambia, por ejemplo, el primer factor ahora explica el 89,11% ($0,9440^2$) y el 90,56% ($0,9516^2$) de la varianza total de la variable $\ln(100 - e_0)$ femenina y masculina, respectivamente.

Para el caso de la mortalidad infantil femenina y masculina, la rotación hace que el primer factor explique el 82,52% ($0,9084^2$) y el 81,10% ($0,9006^2$) de su varianza total en forma respectiva. También, permite encontrar mayor participación del factor 2 en la explicación de la varianza de la mortalidad en grupos de menor edad de ambos sexos respecto de los porcentajes encontrados por Ledermann y Breas (1959).

Tasas de mortalidad, esperanza de vida y medidas de precisión estimadas

Considerando las cargas factoriales rotadas, es posible reconstruir las 111 tablas de vida utilizadas en el estudio de modo que para el año j , la esperanza de vida al nacer y la tasa de mortalidad del grupo de edad x puede estimarse a partir de estos coeficientes de la siguiente forma (se ilustra para las primeras tres ecuaciones de cada sexo).

Sexo femenino

$$\ln(100 - e_0): \hat{z}_{1j} = 0,9440F_{1j} + 0,3211F_{2j} - 0,0259F_{3j}$$

$$\ln(m(0,1)): \hat{z}_{2j} = 0,9084F_{1j} + 0,3728F_{2j} + 0,0061F_{3j}$$

$$\ln(m(1,4)): \hat{z}_{3j} = 0,8824F_{1j} + 0,4164F_{2j} + 0,0420F_{3j}$$

Sexo Masculino

$$\ln(100 - e_0): \hat{z}_{21j} = 0,9516F_{1j} + 0,3005F_{2j} - 0,0410F_{3j}$$

$$\ln(m(0,1)): \hat{z}_{22j} = 0,9006F_{1j} + 0,3853F_{2j} - 0,0013F_{3j}$$

$$\ln(m(1,4)): \hat{z}_{23j} = 0,8920F_{1j} + 0,4064F_{2j} + 0,0300F_{3j}$$

Recordando que $\hat{x}_{ij} = \hat{z}_{ij}\sigma_i + \mu_i$, los nuevos coeficientes para las ecuaciones de las variables x , es decir, para los logaritmos se obtiene de las siguientes relaciones

Sexo femenino

$$\ln(100 - e_0) = \hat{x}_{1j} = (0,9440F_{1j} + 0,3211F_{2j} - 0,0259F_{3j})\sigma_1 + \mu_1$$

$$\ln(m(0,1)) = \hat{x}_{2j} = (0,9084F_{1j} + 0,3728F_{2j} + 0,0061F_{3j})\sigma_2 + \mu_2$$

$$\ln(m(1,4)) = \hat{x}_{3j} = (0,8824F_{1j} + 0,4164F_{2j} + 0,0420F_{3j})\sigma_3 + \mu_3$$

Sexo Masculino

$$\ln(100 - e_0) = \hat{x}_{21j} = (0,9516F_{1j} + 0,3005F_{2j} - 0,0410F_{3j})\sigma_{21} + \mu_{21}$$

$$\ln(m(0,1)) = \hat{x}_{22j} = (0,9006F_{1j} + 0,3853F_{2j} - 0,0013F_{3j})\sigma_{22} + \mu_{22}$$

$$\ln(m(1,4)) = \hat{x}_{23j} = (0,8920F_{1j} + 0,4064F_{2j} + 0,0300F_{3j})\sigma_{22} + \mu_{22}$$

Los coeficientes de las ecuaciones anteriores pueden observarse en el cuadro 7. La reproducción de las tasas de mortalidad y la esperanza de vida al nacer a partir de los coeficientes A_1 , A_2 y A_3 para el año 2010 muestra algunas estimaciones con diferencias que no alcanzan el 1% respecto de sus valores observados, por ejemplo, en la mortalidad femenina del grupo de 5 a 9 (0,0 %). Si se considera la esperanza de vida al nacer, el sexo femenino presenta una subestimación de 0,52 años, lo que representa un error porcentual del 0,6%, mientras tanto en los hombres se observa una sobrestimación de 0,14 que significa un error del 0,2% (Cuadro 8).

En la mortalidad infantil, los errores porcentuales aumentan considerablemente hasta alcanzar un 25,8% en las mujeres y un 20,3% en los hombres que se ven reflejados en diferencias absolutas de alrededor de 20 individuos por cada 10 000. En los grupos de 1-4 y de 5-9 las estimaciones difieren, según el sexo, pues el error porcentual femenino de este primer grupo es bastante mayor al mostrado en los hombres, mientras que en el grupo siguiente el error en el sexo masculino se eleva con respecto al del caso femenino; no obstante, por las bajas mortalidades en estos grupos las diferencias absolutas no alcanzan la unidad por cada 10 000 habitantes. Otra diferencia porcentual alta (21,8%) se observa en la mortalidad masculina para el grupo de 15 a 19; sin embargo, la diferencia absoluta es alrededor de 2 por cada 10 000 habitantes (Cuadro 8)

En las edades de 25 a 74 años, las diferencias porcentuales varían entre 0,1% y el 17,9%. Los grupos donde estos porcentajes son mayores se encuentran diferencias absolutas no mayores a tres defunciones por cada 10 000 habitantes (Cuadro 8).

En las edades avanzadas de 75 años y más las diferencias porcentuales son alrededor del 8%; sin embargo, llama la atención la subestimación de la tasa masculina del grupo de 85 y más, pues en ella la diferencia absoluta es de 128 muertes por cada 10 000 habitantes.

Considerando la esperanza de vida al nacer para el periodo en estudio, las estimaciones muestran una tendencia similar a la trayectoria de los valores observados con una leve diferencia, en los últimos años, para el sexo femenino, mientras tanto en los hombres el comportamiento tiende a mejorarse, sobre todo, en los últimos 20 años (Gráfica 1).

En términos generales, el ajuste del modelo presenta un error medio de 3,3% y 1,9% en la estimación de la esperanza de vida femenina y masculina, respectivamente. Sin embargo, para la mortalidad infantil el error medio relativo llega a ser alrededor del 28% mientras que en el grupo de 1 a 4 esta medida alcanza el 44,5% en las mujeres y 38,2% en los hombres. Por otro lado, los errores porcentuales promedio son de 10 % o más en las grupos de edad menores a 45 años. Mientras tanto en las edades superiores esta medida no sobrepasa el 10% salvo algunas excepciones. Los errores porcentuales mínimos para las tasas de mortalidad se observan en el grupo de 60 a 64 años (6,2%) en el caso femenino y en el grupo de 70 a 74 años (6,3%) para el sexo masculino (Cuadro 9).

El comportamiento de los errores descrito anteriormente presenta gran similitud con el encontrado por Ledermann y Breas en su estudio de 1959.

Finalmente, la incorporación de las dos componentes adicionales sugeridas por Ledermann y Breas (1959) permite obtener una reducción en el error medio relativo en la mortalidad de los distintos grupos de edad (Cuadro 10), pero con mayor influencia en las edades 0-1 y 1-4. Por ejemplo, en el caso de la mortalidad infantil femenina, este error pasa de 27,5% a 9,3%, es decir disminuye alrededor de un 66%, mientras que en el grupo de 1 a 4 la disminución es de un 73%; para la mortalidad infantil masculina el error medio pasa de 28,0% a 8,1%, lo que representa una reducción del 71%. También es destacable la reducción del error correspondiente en la estimación en la esperanza de vida al nacer de las mujeres (45%) y para el grupo de 5 a 9 años (40% en las mujeres y 50% en los hombres). Por otro lado, para la mayoría de las edades de 35 años o más la reducción del error medio es más fuerte en el sexo masculino y en algunos grupos las diferencias entre hombres y mujeres tienden a disminuir.

4. DISCUSION

El empleo de la técnica de análisis factorial ha permitido identificar tres factores principales que explican en un 96,6% la variabilidad de la mortalidad de Costa Rica. Por otro lado, la incorporación de dos componentes adicionales logra reducir de manera importante el error relativo medio en la mortalidad de los grupos de menor edad (0-4); además, muestra que la reducción de este error es más fuerte en el sexo masculino para las edades de 35 años y más.

Los resultados obtenidos en el análisis factorial concuerdan con lo expuesto por Ledermann y Breas (1959); sin embargo, a diferencia de la investigación original, el análisis de este estudio modela la evolución histórica de un solo país, y no la experiencia compartida entre países.

Además, indicar que los resultados del análisis factorial se enfocan en determinar la cantidad de factores influyentes en la mortalidad y, en cierta medida, analizar el comportamiento de los mismos a modo que puedan identificarse; sin embargo, no permiten mostrar otros aspectos como, por ejemplo, las relaciones que asocian estos componentes con los factores causales presuntos de la mortalidad como lo sugiere Naciones Unidas (1962).

Una limitación que puede señalarse en el desarrollo de la presente investigación, es la conformación de la base de datos original, la cual está compuesta por dos fuentes distintas, la primera fue proporcionada por el historiador Héctor Pérez Brignoli y comprende información de 1900 a 1949; la segunda fuente proviene de las tablas de mortalidad elaboradas en conjunto por el CCP-INEC y que contiene los datos referentes al periodo 1950-2010. Al contar con fuentes de información distintas, se corre el riesgo de encontrar incompatibilidades entre ellas (González y Guerrero, 2007) como las halladas en la tasa de mortalidad del grupo abierto de 85 y más, la cual, presenta un comportamiento más irregular con un salto importante en 1950. Este salto podría ser explicado por el cambio en el método con que se calculó. Antes de 1950 la tasa se obtiene por el cociente entre las defunciones totales de personas con edad de 85 años y más y el total de la población de dicho grupo ($D(85 \text{ más}) / N(85 \text{ más})$); es decir, con valores observados en la población, mientras que de 1950 en adelante se estimó como $d(85 \text{ más}) / T85$, esto es, con valores teóricos de la tabla de vida. La estructura por edades de la tabla de vida, que corresponde a una población estacionaria es, sustancialmente, más envejecida que la estructura de la población real, como consecuencia la $m(85 \text{ más})$ de la tabla va a ser sustancialmente mayor que la $m(85 \text{ más})$ de la población.

Finalmente, se espera que esta serie de estimaciones contribuyan al estudio del comportamiento de la mortalidad y de sus implicaciones para el desarrollo del país y, a su vez, sirva de base para que futuras investigaciones complementen los alcances de la presente investigación.

5. REFERENCIAS

- Bongaarts, J. (2005). Long-range trends in adult mortality: Models and projection methods. *Demography*, 42 (1), 23-49.
- CCP / INEC. (2008). *Costa Rica: Estimaciones y Proyecciones de Población por sexo y edad (cifras actualizadas) 1950-2100*. San José, Costa Rica: Publicaciones del Instituto Nacional de Estadística y Censos (INEC).
- González, C. y Guerrero, V. (2007). *Pronósticos estadísticos de mortalidad y su impacto sobre el Sistema de Pensiones de México*. Consulta del 25 de enero del 2011 de http://www.consar.gob.mx/premio_pensiones/pdf/2007
- Hernández, O. (1998). *Temas de análisis estadístico multivariado*. San José, Costa Rica. Editorial de la Universidad de Costa Rica.
- Ledermann, S. y Breas, J. (1959). Les dimensions de la mortalité. *Population*, 637-682.

- Naciones Unidas. (1962). Análisis factorial de las tasas de mortalidad por edad y por sexo. *Boletín de Población de las Naciones Unidas*, 6, 155-209.
- Pardo, A. y Ruiz, M. (2002). *SPPS 11. Guía para el análisis de datos*. (pp. 410-411). Madrid, España: McGraw-Hill.
- Peláez, E. (1998). Análisis del nivel y cambio de la mortalidad joven y adulta en la provincia de Córdoba entre 1980 y 1995. *XI Encontro Nacional de Estudos Populacionais da ABEP*. Consulta del 22 de octubre del 2010 de <http://www.abep.nepo.unicamp.br/docs/anais/PDF/1998/a214.pdf>.
- Pérez, H. (2010). *La población de Costa Rica 1750-200: Una historia experimental*. Costa Rica: Editorial UCR.

Cuadro 1. Costa Rica. Principales medidas descriptivas para cada variable en estudio por sexo. Periodo 1900-2010.

Variable (logaritmo)	Femenino			Masculino		
	No.	Promedio	Desviación Estándar	No.	Promedio	Desviación Estándar
100 - e_0	1	3,60262	0,45813	21	3,70373	0,38687
$m(0-1)$	2	-2,67348	1,28260	22	-2,51083	1,22544
$m(1-4)$	3	-5,15610	1,71343	23	-5,06209	1,65520
$m(5-9)$	4	-6,48403	1,46964	24	-6,35191	1,35337
$m(10-14)$	5	-6,92475	1,15698	25	-6,70895	0,99550
$m(15-19)$	6	-6,47401	1,17788	26	-6,16706	0,83004
$m(20-24)^a$	7	-6,11098	1,31621	27	-5,71144	0,89599
$m(25-29)^a$	8	-5,95650	1,31191	28	-5,60441	0,90439
$m(30-34)$	9	-5,74279	1,22165	29	-5,46392	0,89365
$m(35-39)$	10	-5,44874	1,19905	30	-5,24442	0,88170
$m(40-44)$	11	-5,20933	1,06310	31	-4,98948	0,83355
$m(45-49)$	12	-4,99176	0,93423	32	-4,73536	0,75468
$m(50-54)$	13	-4,73393	0,74912	33	-4,43869	0,65669
$m(55-59)$	14	-4,37308	0,70150	34	-4,11578	0,58056
$m(60-64)$	15	-3,96692	0,61302	35	-3,73168	0,51500
$m(65-69)$	16	-3,63371	0,49160	36	-3,41792	0,37622
$m(70-74)$	17	-3,22710	0,42719	37	-3,02856	0,31793
$m(75-79)$	18	-2,97182	0,30174	38	-2,70929	0,25918
$m(80-84)$	19	-2,53625	0,21393	39	-2,37580	0,17059
$m(85 \text{ más})$	20	-2,11133	0,30638	40	-1,98050	0,24892

^a En el caso del sexo femenino, las tasas de estos dos grupos para el año 2003 fueron calculadas nuevamente a partir de los datos de población y defunciones.

Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Cuadro 2. Costa Rica. Coeficientes de correlación entre las variables del sexo femenino. Período 1900-2010.

Variable (logaritmo)	100 - e ₀	m (0-1)	m (1-4)	m (5-9)	m (10-14)	m (15-19)	m (20-24)	m (25-29)	m (30-34)	m (35-39)	m (40-44)	m (45-49)	m (50-54)	m (55-59)	m (60-64)	m (65-69)	m (70-74)	m (75-79)	m (80-84)	m (85 más)
100 - e ₀	1,000	0,988	0,978	0,991	0,989	0,988	0,990	0,991	0,993	0,993	0,994	0,991	0,993	0,980	0,991	0,976	0,951	0,534	0,593	-0,564
m(0-1)		1,000	0,995	0,990	0,970	0,963	0,968	0,971	0,973	0,977	0,973	0,965	0,975	0,955	0,976	0,969	0,951	0,575	0,617	-0,533
m(1-4)			1,000	0,986	0,960	0,950	0,952	0,956	0,958	0,964	0,958	0,949	0,962	0,941	0,967	0,968	0,954	0,606	0,643	-0,508
m(5-9)				1,000	0,980	0,978	0,979	0,980	0,981	0,983	0,980	0,973	0,981	0,963	0,983	0,970	0,950	0,576	0,597	-0,577
m(10-14)					1,000	0,982	0,982	0,984	0,983	0,981	0,986	0,982	0,981	0,965	0,978	0,960	0,933	0,534	0,553	-0,579
m(15-19)						1,000	0,992	0,991	0,989	0,988	0,987	0,986	0,987	0,977	0,981	0,966	0,929	0,514	0,549	-0,616
m(20-24)							1,000	0,995	0,994	0,993	0,991	0,989	0,990	0,977	0,982	0,959	0,928	0,502	0,540	-0,621
m(25-29)								1,000	0,996	0,994	0,992	0,991	0,990	0,979	0,984	0,963	0,929	0,504	0,541	-0,614
m(30-34)									1,000	0,996	0,994	0,993	0,991	0,980	0,984	0,962	0,929	0,503	0,542	-0,608
m(35-39)										1,000	0,992	0,989	0,991	0,979	0,986	0,968	0,934	0,522	0,552	-0,609
m(40-44)											1,000	0,994	0,991	0,976	0,985	0,961	0,939	0,521	0,556	-0,592
m(45-49)												1,000	0,992	0,982	0,984	0,959	0,934	0,478	0,570	-0,570
m(50-54)													1,000	0,984	0,991	0,974	0,947	0,526	0,594	-0,575
m(55-59)														1,000	0,982	0,965	0,941	0,430	0,620	-0,525
m(60-64)															1,000	0,984	0,961	0,555	0,631	-0,560
m(65-69)																1,000	0,961	0,588	0,670	-0,498
m(70-74)																	1,000	0,587	0,760	-0,444
m(75-79)																		1,000	0,346	-0,442
m(80-84)																			1,000	0,098
m(85 más)																				1,000

Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Cuadro 4. Costa Rica. Coeficientes de correlación entre las variables de cada sexo. Periodo 1900-2010.

	Masculino																			
	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>	<i>m</i>			
(logaritmo) 100- <i>e</i> ₀	(0-1)	(1-4)	(5-9)	(10-14)	(15-19)	(20-24)	(25-29)	(30-34)	(35-39)	(40-44)	(45-49)	(50-54)	(55-59)	(60-64)	(65-69)	(70-74)	(75-79)	(80-84)	(85 más)	
100- <i>e</i> ₀	0,999	0,985	0,982	0,988	0,988	0,969	0,964	0,957	0,965	0,973	0,975	0,979	0,978	0,970	0,983	0,965	0,952	0,669	0,413	-0,631
<i>m</i> (0-1)	0,981	0,999	0,995	0,989	0,978	0,932	0,924	0,916	0,924	0,937	0,938	0,944	0,943	0,936	0,956	0,947	0,938	0,681	0,443	-0,595
<i>m</i> (1-4)	0,969	0,995	0,998	0,987	0,971	0,913	0,898	0,889	0,899	0,914	0,916	0,923	0,923	0,918	0,943	0,940	0,940	0,707	0,480	-0,554
<i>m</i> (5-9)	0,986	0,989	0,990	0,993	0,986	0,947	0,935	0,926	0,936	0,947	0,949	0,956	0,955	0,946	0,965	0,959	0,949	0,696	0,439	-0,621
<i>m</i> (10-14)	0,990	0,966	0,965	0,979	0,985	0,970	0,963	0,954	0,959	0,966	0,975	0,973	0,972	0,958	0,976	0,956	0,939	0,675	0,400	-0,633
<i>m</i> (15-19)	0,991	0,959	0,959	0,972	0,980	0,982	0,971	0,960	0,971	0,972	0,977	0,980	0,979	0,975	0,978	0,966	0,943	0,676	0,387	-0,648
<i>m</i> (20-24)	0,993	0,965	0,960	0,974	0,977	0,975	0,976	0,969	0,978	0,982	0,980	0,985	0,983	0,971	0,976	0,957	0,937	0,649	0,373	-0,673
<i>m</i> (25-29)	0,993	0,968	0,965	0,978	0,980	0,974	0,973	0,966	0,974	0,980	0,979	0,983	0,980	0,972	0,978	0,962	0,939	0,647	0,367	-0,676
<i>m</i> (30-34)	0,994	0,970	0,966	0,977	0,977	0,972	0,973	0,966	0,976	0,983	0,979	0,985	0,983	0,972	0,977	0,960	0,936	0,640	0,361	-0,680
<i>m</i> (35-39)	0,993	0,975	0,971	0,981	0,979	0,967	0,964	0,956	0,968	0,976	0,972	0,979	0,977	0,970	0,977	0,967	0,945	0,661	0,385	-0,667
<i>m</i> (40-44)	0,995	0,970	0,963	0,977	0,983	0,973	0,975	0,969	0,976	0,983	0,984	0,987	0,985	0,968	0,983	0,959	0,944	0,649	0,393	-0,659
<i>m</i> (45-49)	0,994	0,961	0,956	0,967	0,978	0,978	0,980	0,978	0,982	0,988	0,987	0,991	0,989	0,975	0,983	0,953	0,938	0,624	0,379	-0,650
<i>m</i> (50-54)	0,993	0,972	0,968	0,978	0,981	0,970	0,968	0,962	0,970	0,977	0,978	0,983	0,986	0,976	0,986	0,967	0,954	0,666	0,417	-0,631
<i>m</i> (55-59)	0,982	0,951	0,950	0,955	0,968	0,971	0,964	0,963	0,969	0,973	0,969	0,974	0,976	0,985	0,975	0,954	0,939	0,601	0,384	-0,612
<i>m</i> (60-64)	0,989	0,973	0,973	0,979	0,982	0,961	0,953	0,944	0,953	0,962	0,966	0,971	0,975	0,971	0,987	0,976	0,969	0,697	0,469	-0,598
<i>m</i> (65-69)	0,971	0,966	0,972	0,969	0,969	0,943	0,923	0,908	0,919	0,928	0,939	0,943	0,949	0,962	0,973	0,978	0,967	0,747	0,507	-0,510
<i>m</i> (70-74)	0,945	0,949	0,956	0,948	0,952	0,909	0,894	0,888	0,891	0,898	0,907	0,917	0,915	0,920	0,946	0,925	0,964	0,711	0,616	-0,435
<i>m</i> (75-79)	0,514	0,587	0,594	0,608	0,540	0,417	0,394	0,343	0,376	0,399	0,441	0,441	0,464	0,418	0,531	0,611	0,640	0,914	0,626	-0,160
<i>m</i> (80-84)	0,578	0,614	0,639	0,584	0,610	0,557	0,517	0,531	0,513	0,518	0,530	0,544	0,544	0,592	0,606	0,574	0,671	0,467	0,770	0,114
<i>m</i> (85 más)	-0,576	-0,539	-0,527	-0,587	-0,553	-0,561	-0,578	-0,536	-0,567	-0,567	-0,569	-0,576	-0,576	-0,521	-0,553	-0,572	-0,531	-0,474	-0,083	0,819

Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Cuadro 5. Costa Rica. Valores residuales entre las correlaciones observadas y estimadas para cada par de variables obtenidos por los tres primeros factores por sexo. Periodo 1900-2010

Variable	Femenino						Masculino							
	$100 - e_0$	$m(0-1)$	$m(1-4)$	$m(5-9)$	$m(75-79)$	$m(80-84)$	$m(85 \text{ más})$	$100 - e_0$	$m(0-1)$	$m(1-4)$	$m(5-9)$	$m(75-79)$	$m(80-84)$	$m(85 \text{ más})$
Femenino														
$100 - e_0$	-	0,011	0,013	0,007	0,006	-0,002	0,013	0,003	0,011	0,011	0,007	-0,004	-0,014	-0,002
$m(0-1)$		-	0,038	0,020	0,009	0,003	0,023	0,005	0,037	0,033	0,021	-0,022	-0,028	-0,014
$m(1-4)$			-	0,024	0,007	0,004	0,026	0,006	0,040	0,040	0,026	-0,024	-0,033	-0,017
$m(5-9)$				-	0,001	0,006	0,010	0,002	0,021	0,022	0,016	-0,014	-0,017	-0,011
$m(75-79)$					-	-0,025	0,054	0,002	0,011	0,002	0,007	0,005	-0,037	0,000
$m(80-84)$						-	-0,048	-0,004	0,006	0,006	0,004	-0,035	0,046	0,000
$m(85 \text{ más})$							-	0,007	0,023	0,017	0,017	0,028	-0,104	0,031
Masculino														
$100 - e_0$								-	0,004	0,005	0,003	0,001	-0,006	0,002
$m(0-1)$									-	0,035	0,023	-0,027	-0,030	-0,017
$m(1-4)$										-	0,023	-0,020	-0,029	-0,013
$m(5-9)$											-	-0,017	-0,023	-0,011
$m(75-79)$												-	-0,034	0,048
$m(80-84)$													-	-0,031
$m(85 \text{ más})$														-

Fuente: Elaboración propia a partir del los datos de INEC-CCP y Héctor Pérez

Cuadro 6. Costa Rica. Coeficientes de correlación entre las variables y los tres primeros componentes principales (\hat{a}_k) y los tres componentes factoriales rotados (\hat{A}_k) por sexo.

Periodo 1900-2010.							
Variable (logaritmo)	\hat{a}_1	\hat{a}_2	\hat{a}_3	\hat{A}_1	\hat{A}_2	\hat{A}_3	Comunalidad $\sum \hat{a}^2 = \sum \hat{A}^2$
Femenino							
100 - e_0	0,9972	-0,0121	0,0178	0,9440	0,3211	-0,0259	0,9949
m(0-1)	0,9805	0,0526	-0,0043	0,9084	0,3728	0,0061	0,9642
m(1-4)	0,9699	0,1128	-0,0157	0,8824	0,4164	0,0420	0,9538
m(5-9)	0,9879	0,0289	-0,0337	0,9127	0,3795	-0,0316	0,9780
m(10-14)	0,9880	-0,0306	-0,0083	0,9324	0,3233	-0,0570	0,9772
m(15-19)	0,9915	-0,0554	-0,0132	0,9397	0,3116	-0,0784	0,9863
m(20-24)	0,9924	-0,0815	-0,0069	0,9479	0,2905	-0,0930	0,9916
m(25-29)	0,9933	-0,0811	-0,0051	0,9491	0,2898	-0,0914	0,9932
m(30-34)	0,9934	-0,0858	0,0005	0,9517	0,2830	-0,0910	0,9942
m(35-39)	0,9937	-0,0571	-0,0152	0,9417	0,3126	-0,0811	0,9910
m(40-44)	0,9945	-0,0564	0,0059	0,9479	0,2991	-0,0660	0,9923
m(45-49)	0,9918	-0,0710	0,0557	0,9618	0,2551	-0,0421	0,9919
m(50-54)	0,9955	-0,0183	0,0213	0,9447	0,3141	-0,0278	0,9919
m(55-59)	0,9818	-0,0455	0,1234	0,9648	0,2229	0,0234	0,9812
m(60-64)	0,9942	0,0436	0,0133	0,9279	0,3597	0,0113	0,9905
m(65-69)	0,9789	0,1353	0,0202	0,8955	0,4101	0,0828	0,9769
m(70-74)	0,9547	0,2322	0,0819	0,8683	0,4240	0,1961	0,9722
m(75-79)	0,5468	0,5165	-0,6075	0,2400	0,9345	-0,0629	0,9348
m(80-84)	0,5960	0,6058	0,4417	0,5463	0,3049	0,7252	0,9173
m(85 más)	-0,5843	0,3734	0,6080	-0,4677	-0,3631	0,7071	0,8506
Masculino							
100 - e_0	0,9978	-0,0382	0,0233	0,9516	0,3005	-0,0410	0,9976
m(0-1)	0,9776	0,0572	-0,0199	0,9006	0,3853	-0,0013	0,9594
m(1-4)	0,9759	0,0955	-0,0147	0,8920	0,4064	0,0300	0,9618
m(5-9)	0,9852	0,0386	-0,0703	0,8982	0,4095	-0,0498	0,9770
m(10-14)	0,9885	0,0218	0,0206	0,9292	0,3386	0,0008	0,9780
m(15-19)	0,9748	-0,0933	0,0939	0,9608	0,2090	-0,0313	0,9678
m(20-24)	0,9690	-0,1540	0,0937	0,9685	0,1673	-0,0751	0,9716
m(25-29)	0,9599	-0,1695	0,1619	0,9814	0,1081	-0,0389	0,9763
m(30-34)	0,9682	-0,1742	0,1121	0,9770	0,1414	-0,0769	0,9803
m(35-39)	0,9749	-0,1629	0,0993	0,9774	0,1597	-0,0778	0,9869
m(40-44)	0,9791	-0,1170	0,0691	0,9633	0,2115	-0,0656	0,9771
m(45-49)	0,9834	-0,1093	0,0709	0,9662	0,2169	-0,0589	0,9840
m(50-54)	0,9837	-0,0940	0,0523	0,9582	0,2396	-0,0608	0,9792
m(55-59)	0,9751	-0,0636	0,1210	0,9619	0,2104	0,0089	0,9695
m(60-64)	0,9888	0,0139	0,0307	0,9339	0,3267	0,0021	0,9789
m(65-69)	0,9735	0,0636	-0,0772	0,8800	0,4265	-0,0362	0,9577
m(70-74)	0,9629	0,1836	-0,0458	0,8524	0,4807	0,0726	0,9630
m(75-79)	0,6892	0,4874	-0,4625	0,4186	0,8666	0,0121	0,9264
m(80-84)	0,4373	0,7866	-0,0291	0,2327	0,6858	0,5351	0,8109
m(85 más)	-0,6274	0,6670	0,2836	-0,6581	0,0330	0,6962	0,9189
$\sum \hat{a}_{ik}^2$	34,8660	2,4138	1,3664	30,9678	5,7364	1,9420	38,6462
Contribución	87,1651	6,0345	3,4160	77,4195	14,3409	4,8550	96,6155

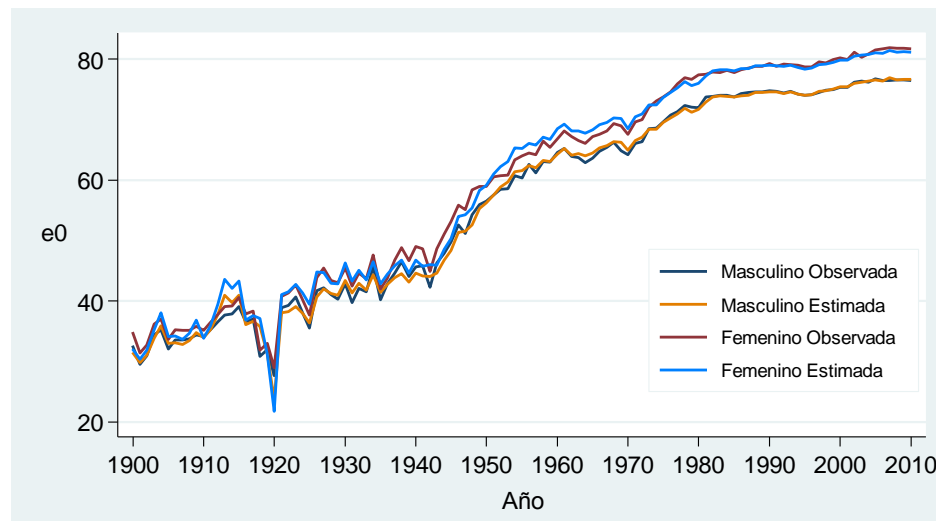
Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Cuadro 7. Costa Rica. Coeficientes de la ecuación de estimación de las variables para los tres primeros factores por sexo. Periodo 1900-2010.

Logaritmo	No.	F_1	F_2	F_3	μ
Femenino					
$100 - e_0$	1	0,4325	0,1471	-0,0119	3,6026
$m(0-1)$	2	0,4162	0,1708	0,0028	-2,6735
$m(1-4)$	3	0,4043	0,1908	0,0192	-5,1561
$m(5-9)$	4	0,4181	0,1739	-0,0145	-6,4840
$m(10-14)$	5	0,4272	0,1481	-0,0261	-6,9247
$m(15-19)$	6	0,4305	0,1428	-0,0359	-6,4740
$m(20-24)$	7	0,4343	0,1331	-0,0426	-6,1110
$m(25-29)$	8	0,4348	0,1328	-0,0419	-5,9565
$m(30-34)$	9	0,4360	0,1297	-0,0417	-5,7428
$m(35-39)$	10	0,4314	0,1432	-0,0371	-5,4487
$m(40-44)$	11	0,4342	0,1370	-0,0302	-5,2093
$m(45-49)$	12	0,4406	0,1169	-0,0193	-4,9918
$m(50-54)$	13	0,4328	0,1439	-0,0128	-4,7339
$m(55-59)$	14	0,4420	0,1021	0,0107	-4,3731
$m(60-64)$	15	0,4251	0,1648	0,0052	-3,9669
$m(65-69)$	16	0,4102	0,1879	0,0379	-3,6337
$m(70-74)$	17	0,3978	0,1942	0,0898	-3,2271
$m(75-79)$	18	0,1099	0,4281	-0,0288	-2,9718
$m(80-84)$	19	0,2503	0,1397	0,3322	-2,5363
$m(85 \text{ más})$	20	-0,2142	-0,1664	0,3240	-2,1113
Masculino					
$100 - e_0$	21	0,4360	0,1377	-0,0188	3,7037
$m(0-1)$	22	0,4126	0,1765	-0,0006	-2,5108
$m(1-4)$	23	0,4087	0,1862	0,0138	-5,0621
$m(5-9)$	24	0,4115	0,1876	-0,0228	-6,3519
$m(10-14)$	25	0,4257	0,1551	0,0003	-6,7090
$m(15-19)$	26	0,4402	0,0957	-0,0143	-6,1671
$m(20-24)$	27	0,4437	0,0766	-0,0344	-5,7114
$m(25-29)$	28	0,4496	0,0495	-0,0178	-5,6044
$m(30-34)$	29	0,4476	0,0648	-0,0352	-5,4639
$m(35-39)$	30	0,4478	0,0732	-0,0357	-5,2444
$m(40-44)$	31	0,4413	0,0969	-0,0301	-4,9895
$m(45-49)$	32	0,4426	0,0994	-0,0270	-4,7354
$m(50-54)$	33	0,4390	0,1097	-0,0278	-4,4387
$m(55-59)$	34	0,4407	0,0964	0,0041	-4,1158
$m(60-64)$	35	0,4278	0,1497	0,0010	-3,7317
$m(65-69)$	36	0,4032	0,1954	-0,0166	-3,4179
$m(70-74)$	37	0,3905	0,2202	0,0332	-3,0286
$m(75-79)$	38	0,1918	0,3970	0,0055	-2,7093
$m(80-84)$	39	0,1066	0,3142	0,2452	-2,3758
$m(85 \text{ más})$	40	-0,3015	0,0151	0,3190	-1,9805

Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Gráfica 1. Costa Rica. Esperanza de vida al nacer según año por sexo. Periodo 1900-2010.



Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Cuadro 8. Costa Rica. Tasas de mortalidad por grupo de edad y esperanza de vida al nacer observada y estimada por sexo. Año 2010.

Variable	Femenino				Masculino			
	Observada	Estimada	Absoluta	Porcentaje	Observada	Estimada	Absoluta	Porcentaje
e_0	81,71	81,19	-0,52	-0,6	76,49	76,63	0,14	0,2
$m(0-1)$	79,00	99,37	20,37	25,8	105,80	127,29	21,49	20,3
$m(1-4)$	3,20	3,93	0,73	22,8	4,90	4,82	-0,08	-1,6
$m(5-9)$	1,70	1,70	0,00	0,0	1,90	2,29	0,39	20,4
$m(10-14)$	2,00	1,90	-0,10	-5,2	2,60	2,78	0,18	6,9
$m(15-19)$	2,90	2,97	0,07	2,5	8,90	6,96	-1,94	-21,8
$m(20-24)$	4,50	3,66	-0,84	-18,6	13,10	10,82	-2,28	-17,4
$m(25-29)$	4,60	4,29	-0,31	-6,7	14,30	12,26	-2,04	-14,3
$m(30-34)$	6,40	6,05	-0,35	-5,4	17,30	14,21	-3,09	-17,9
$m(35-39)$	8,70	8,03	-0,67	-7,7	20,50	17,60	-2,91	-14,2
$m(40-44)$	11,90	12,32	0,42	3,5	24,40	22,92	-1,49	-6,1
$m(45-49)$	18,20	18,71	0,51	2,8	30,10	32,37	2,27	7,5
$m(50-54)$	28,00	29,70	1,70	6,1	48,40	48,97	0,57	1,2
$m(55-59)$	47,70	47,36	-0,34	-0,7	72,10	73,87	1,77	2,4
$m(60-64)$	69,90	74,54	4,64	6,6	112,00	111,93	-0,07	-0,1
$m(65-69)$	113,20	119,87	6,67	5,9	180,60	185,50	4,90	2,7
$m(70-74)$	193,80	192,90	-0,90	-0,5	275,60	287,13	11,53	4,2
$m(75-79)$	318,60	333,44	14,84	4,7	478,50	440,03	-38,47	-8,0
$m(80-84)$	532,50	561,43	28,93	5,4	734,90	709,12	-25,78	-3,5
$m(85 \text{ más})$	1 331,00	1374,50	43,50	3,3	1 552,80	1 424,78	-128,03	-8,2

Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Cuadro 9. Costa Rica. Error típico en la estimación de las variables a partir de los tres primeros factores por sexo. Periodo 1900-2010.

Variable (logaritmo)	σ	r^2	s^2	s	S	$\exp(S)$	Error %
Femenino							
$100 - e_0$	0,4581	0,9949	0,0051	0,0711	0,0326	1,0331	3,3
$m(0-1)$	1,2826	0,9642	0,0358	0,1891	0,2426	1,2745	27,5
$m(1-4)$	1,7134	0,9538	0,0462	0,2150	0,3684	1,4454	44,5
$m(5-9)$	1,4696	0,9780	0,0220	0,1483	0,2180	1,2436	24,4
$m(10-14)$	1,1570	0,9772	0,0228	0,1510	0,1748	1,1910	19,1
$m(15-19)$	1,1779	0,9863	0,0137	0,1171	0,1379	1,1478	14,8
$m(20-24)$	1,3162	0,9916	0,0084	0,0914	0,1204	1,1279	12,8
$m(25-29)$	1,3119	0,9932	0,0068	0,0824	0,1082	1,1142	11,4
$m(30-34)$	1,2217	0,9942	0,0058	0,0764	0,0933	1,0978	9,8
$m(35-39)$	1,1990	0,9910	0,0090	0,0948	0,1137	1,1204	12,0
$m(40-44)$	1,0631	0,9923	0,0077	0,0880	0,0936	1,0981	9,8
$m(45-49)$	0,9342	0,9919	0,0081	0,0900	0,0840	1,0877	8,8
$m(50-54)$	0,7491	0,9919	0,0081	0,0901	0,0675	1,0698	7,0
$m(55-59)$	0,7015	0,9812	0,0188	0,1373	0,0963	1,1011	10,1
$m(60-64)$	0,6130	0,9905	0,0095	0,0977	0,0599	1,0617	6,2
$m(65-69)$	0,4916	0,9769	0,0231	0,1520	0,0747	1,0776	7,8
$m(70-74)$	0,4272	0,9722	0,0278	0,1669	0,0713	1,0739	7,4
$m(75-79)$	0,3017	0,9348	0,0652	0,2553	0,0770	1,0801	8,0
$m(80-84)$	0,2139	0,9173	0,0827	0,2876	0,0615	1,0635	6,3
$m(85 \text{ más})$	0,3064	0,8506	0,1494	0,3865	0,1184	1,1257	12,6
Masculino							
$100 - e_0$	0,3869	0,9976	0,0024	0,0494	0,0191	1,0193	1,9
$m(0-1)$	1,2254	0,9594	0,0406	0,2014	0,2468	1,2800	28,0
$m(1-4)$	1,6552	0,9618	0,0382	0,1955	0,3235	1,3820	38,2
$m(5-9)$	1,3534	0,9770	0,0230	0,1516	0,2051	1,2277	22,8
$m(10-14)$	0,9955	0,9780	0,0220	0,1482	0,1475	1,1590	15,9
$m(15-19)$	0,8300	0,9678	0,0322	0,1794	0,1489	1,1605	16,1
$m(20-24)$	0,8960	0,9716	0,0284	0,1687	0,1511	1,1631	16,3
$m(25-29)$	0,9044	0,9763	0,0237	0,1539	0,1392	1,1494	14,9
$m(30-34)$	0,8937	0,9803	0,0197	0,1402	0,1253	1,1335	13,3
$m(35-39)$	0,8817	0,9869	0,0131	0,1147	0,1011	1,1064	10,6
$m(40-44)$	0,8336	0,9771	0,0229	0,1514	0,1262	1,1345	13,4
$m(45-49)$	0,7547	0,9840	0,0160	0,1265	0,0954	1,1001	10,0
$m(50-54)$	0,6567	0,9792	0,0208	0,1441	0,0946	1,0993	9,9
$m(55-59)$	0,5806	0,9695	0,0305	0,1745	0,1013	1,1066	10,7
$m(60-64)$	0,5150	0,9789	0,0211	0,1453	0,0748	1,0777	7,8
$m(65-69)$	0,3762	0,9577	0,0423	0,2058	0,0774	1,0805	8,0
$m(70-74)$	0,3179	0,9630	0,0370	0,1924	0,0612	1,0631	6,3
$m(75-79)$	0,2592	0,9264	0,0736	0,2712	0,0703	1,0728	7,3
$m(80-84)$	0,1706	0,8109	0,1891	0,4349	0,0742	1,0770	7,7
$m(85 \text{ más})$	0,2489	0,9189	0,0811	0,2847	0,0709	1,0734	7,3

Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez

Cuadro 10. Costa Rica. Error medio relativo en la estimación de las variables con los tres y cinco primeros factores por sexo. Periodo 1900-2010.

Variable	Tres componentes		Cinco componentes	
	Femenino	Masculino	Femenino	Masculino
$100 - e_0$	3,3	1,9	1,8	1,7
$m(0-1)$	27,5	28,0	9,3	8,1
$m(1-4)$	44,5	38,2	12,1	13,0
$m(5-9)$	24,4	22,8	14,5	11,3
$m(10-14)$	19,1	15,9	18,3	14,9
$m(15-19)$	14,8	16,1	13,9	12,0
$m(20-24)$	12,8	16,3	12,7	10,7
$m(25-29)$	11,4	14,9	11,4	10,4
$m(30-34)$	9,8	13,3	9,3	9,4
$m(35-39)$	12,0	10,6	11,2	9,0
$m(40-44)$	9,8	13,4	9,7	9,3
$m(45-49)$	8,8	10,0	8,6	7,1
$m(50-54)$	7,0	9,9	7,0	7,0
$m(55-59)$	10,1	10,7	10,1	8,6
$m(60-64)$	6,2	7,8	6,1	6,3
$m(65-69)$	7,8	8,0	7,0	6,9
$m(70-74)$	7,4	6,3	6,2	5,9
$m(75-79)$	8,0	7,3	6,8	4,1
$m(80-84)$	6,3	7,7	4,8	4,8
$m(85 \text{ más})$	12,6	7,3	7,1	5,3

Fuente: Elaboración propia a partir de los datos de INEC-CCP y Héctor Pérez