

Estimación de una tabla de vida para el Sistema Nacional de pensiones Peruano

*José Artemio Valderrama Torres*¹

Resumen: Se estima una tabla de vida estática empleando el estimador no paramétrico propuesto por Kaplan y Meier y la técnica de regresión polinómica local para la graduación empleando más de 8 millones de registros administrativos recolectados durante los años 2013 al 2017. Los resultados revelan que la brecha en expectativa de vida entre hombres y mujeres es menor respecto a lo reportado por la tabla vigente. Así, la diferencia en esperanza de vida medido a los 65 años de edad se reduce de siete a dos años, explicado por una disminución de la expectativa de vida de las mujeres de 2.4 años y un incremento de 2.1 años en el caso de los hombres. Los resultados también revelan que la deuda actuarial neta estaría sobreestimada en $S/2$ mil millones respecto a lo oficialmente estimado, ello es el resultado de dos efectos que se contraponen: un descenso en la estimación de la reserva en el caso de las mujeres y un incremento en el caso de los hombres, siendo el primer efecto el que domina.

Palabras clave: Tabla de mortalidad, Kaplan-Meier, graduación, regresión polinómica local.

Estimation of a life table for the Peruvian national pension system

Abstract: A static life table is estimated using the nonparametric estimator proposed by Kaplan and Meier and the local polynomial regression technique for the graduation using more than 8 million administrative records collected during the years 2013 to 2017. The results reveal that the gap in life expectancy between men and women is lower than that reported by the current table. Thus, the difference in life expectancy measured at 65 years of age is reduced from seven to two years, explained by a decrease in the life expectancy of women of 2.4 years and an increase in life expectancy from 2.1 years in the case of men. The results also reveal that the net actuarial debt would be overestimated in $S/2$ billion compared to the officially estimated, this is the result of two opposing effects: a decrease in the reserve estimate in the case of women and an increase in the case of men, with the first effect dominating.

Keywords: Mortality table, Kaplan-Meier, graduation, local polynomial regression.

Recibido: 26/11/2018. *Aceptado:* 20/12/2018. *Publicado online:* 31/12/2018.

¹PUCP, Facultad de Ciencias Económicas. e-mail: jvalderramat@pucp.pe

1. Introducción

En el Perú la Oficina de Normalización Previsional (ONP) es la institución encargada de administrar al Sistema Nacional de Pensiones (SNP), el mismo que es regido por el Decreto Ley N° 19990. El SNP es un sistema de reparto en el cual los afiliados aportan para financiar la pensión de los jubilados con los que coexisten. Para tener derecho a una pensión vitalicia el futuro jubilado tiene que haber aportado de manera acumulada al menos 20 años, en un monto no menor al 13% de la remuneración mínima vital vigente al momento del aporte.

La deuda o pasivo que el estado asume con cada afiliado se traduce en garantizar el otorgamiento de la pensión de forma vitalicia, esto es, pagar una pensión hasta el momento del fallecimiento. Dado que la fecha de ocurrencia de este siniestro es un evento aleatorio, estimar el dinero que se requiere para financiar el total de la pensión futura es un problema al que se enfrenta cualquier entidad encargada de administrar pensiones o seguros en general. El instrumento en el que se reportan las probabilidades de fallecimiento o sobrevivencia se conoce como tabla de mortalidad o tabla de vida.

Una tabla de vida es el resultado de un modelo matemático que da cuenta de cómo una población hipotética va disminuyendo desde su nacimiento hasta el fallecimiento de la última persona. Las tablas muestran las probabilidades de muerte en los próximos 12 meses para cada x año de edad. Este último indicador es denotado como q_x para cada edad x .

La tabla de vida vigente, conocida como SP2005, fue propuesta en el año 2006 y fue estimada empleando información de los años 1993 al 2005¹. De acuerdo a esta tabla la esperanza de vida de los hombres de 65 años es de 83, mientras que el de las mujeres es de 90 años. Además, de acuerdo a dicha tabla la probabilidad de muerte de los hombres a partir de los 65 años duplica a las mujeres, esta brecha incluso llega a ser el triple en edades tempranas como los 20 años.

Evaluar si esta brecha es realista para los afiliados y pensionistas del SNP es una hipótesis a contrastar. Si por ejemplo las probabilidades de sobrevivencia de las mujeres están sobreestimadas y la de los hombres subestimadas, entonces el dinero necesario para financiar las pensiones vitalicias para las mujeres también estarían sobrestimadas y subestimados en el caso de los hombres.

Otro elemento que justificaría el esfuerzo por la actualización de la tabla es la obsolescencia de la misma, pues, siendo que las sociedades tienden a envejecer a medida que se dan los avances en salud y la población accede a dichos beneficios, es de suponer que las probabilidades de sobrevivencia mejoren con el tiempo. En ese sentido, si la población tiende a vivir más, puede que el valor monetario necesario para financiar las pensiones esté subestimado, pues la población de 65 a más podría estar viviendo en promedio más de lo que el instrumento actual asume.

En suma, la actualización permitiría contar con estimaciones más realistas de lo que el Estado debe provisionar para hacer frente al pago de las pensiones futuras, y hacer diagnósticos apropiados para la elaboración de políticas.

Este documento se divide en seis secciones y un conjunto de conclusiones. En la sección 2 se presenta la metodología; en la parte 3 se desarrolla la estrategia empírica; y se concluye en la parte 4. Se incluye anexos en los que se dan detalles del proceso de estimación y se muestra la tabla estimada.

¹La SP2005 fue instituido mediante Resolución del Ministerio de Economía y Finanzas N° 146-2007-EF/15.

2. Metodología

Una tabla de vida es una estimación de como una población se va reduciendo debido al fallecimiento de sus miembros. Así, cualquier tabla muestra como una cantidad de personas presentes a la edad 0 disminuye hasta su extinción. Por la manera cómo se recolectan los datos las tablas de vida pueden ser longitudinales o transversales.

En el caso de las primeras se sigue al mismo grupo de unidades hasta que el evento aleatorio ocurra (muerte en el caso que se estudia). Como es señalado por Bueno et al. (2008), diseños de este tipo son propios en los ensayos clínicos o en cualquier ejercicio que no requiera seguir por mucho tiempo al colectivo estudiado; los datos transversales o contemporáneos, de otro lado, son todos aquellos con los que se cuenta en un momento específico en el tiempo.

En el caso de la estimación de una tabla de vida, la estrategia de seguir a un grupo de personas en el tiempo es impracticable, pues se tendría que esperar más de 100 años para el uso de estos datos, momento en el cual dicha información sería inoportuna. Típicamente el modelamiento de una tabla se realiza con información básica de la persona como lo son la edad y el sexo; aunque queda claro que la edad es un determinante de la sobrevivencia, y la evidencia es abrumadora en sostener que las mujeres viven más, estas no son los únicos determinantes.

El apego a ciertos hábitos o conductas como fumar, hacer ejercicios, hábitos alimenticios y cuidados de salud también podrían estar influenciando en la expectativa de vida, no obstante esta es información no suele estar disponible, y no representa un problema si es que los datos corresponden a una población homogénea. El problema existiría si estos resultados se extrapolan a otras poblaciones, como los afiliados al Sistema Privado de Pensiones (SPP) o a toda la población del país. Existe evidencia de que colectividades distintas podrían tener probabilidades de fallecimiento también distintas².

En la práctica, casi todas las tablas de vida se construyen estimando en primer lugar la probabilidad de fallecimiento q_x y luego los demás indicadores³. Las probabilidades q_x que los individuos de edad x fallezcan en el transcurso del año se estiman a partir de las frecuencias relativas de los mismos sucesos con información proveniente de registros administrativos.

Dichas estimaciones, conocidas como tasas brutas de mortalidad, podrían tener errores o presentar altibajos atribuibles exclusivamente al azar; por esta razón es que se procede, en una segunda fase, como señalan Gil et al. (1999), a suavizar los datos originales, es decir, a sustituir los valores brutos por nuevos valores que presenten un desarrollo más regular, sin altibajos y saltos injustificables. En la literatura actuarial este procedimiento es conocido como graduación.

De esta manera es que se va configurando la estrategia de estimación. En Charpentier (2015) se señalan los siguientes pasos para la estimación de una tabla de vida: (i) Estimar la tasa cruda o bruta de mortalidad; (ii) graduar o suavizar dicha tasa y; (iii) validar que el modelo y los datos sean consistentes. A esta recomendación se le debe añadir el tratamiento que se le debe dar a la estimación en las colas (edades tempranas y avanzadas), pues típicamente no existe la suficiente cantidad de datos como para tener un estimador razonable en dichos tramos.

En esta investigación la tasa bruta de mortalidad se estima con un enfoque estadístico. La estrategia para calcular las tasas sigue lo sugerido por Kaplan y Meier (1958) la cual consiste en calcular una tasa para cada edad empleando un estimador no paramétrico conocido como ‘producto límite’. En el caso de la graduación se emplea la técnica de ‘regresión polinómica local’

²En Sánchez (2002) por ejemplo, se cita evidencia en la que se muestra diferencias entre las tablas de mortalidad que se utilizan en el sector asegurador (obtenidas con información de la población general) y la medida de mortalidad de un colectivo de ocupados (perfil habitual de los individuos asegurados).

³Algunos otros indicadores son la probabilidad de sobrevivencia p_x y la cantidad de sobrevivientes l_x .

con determinación del ancho de banda de manera óptima. Detalles metodológicos se incluyen en la siguiente sección.

En cuanto a los datos estos corresponden a registros administrativos de la ONP y el Registro Nacional de Identificación y Estado Civil (RENIEC). En el caso de la ONP, estos corresponden a los afiliados activos y pasivos del SNP, fechas en las que realizaron sus aportes o recibieron sus pensiones, según corresponda. La información proveniente de RENIEC corresponde a la fecha de nacimiento, fecha de fallecimiento y sexo. El horizonte temporal de análisis considera casi 5 años, desde el primero de enero de 2013 a noviembre de 2017.

En el caso de los asegurados, la fecha de inicio es el primer aporte registrado en el período de estudio. La fecha de salida según corresponda podría ser fallecimiento o fecha final del período de estudio. En caso la persona pase a ser pensionista, dicho registro se completa con la información del pensionista, esto es, fallecimiento o censura por fecha final de observación.

Análogamente, para el caso de pensionistas que no figuran como asegurados, debido a que son pensionistas por derecho derivado o son pensionistas titulares que obtuvieron el derecho antes de julio de 1999, la fecha de inicio es la primera pensión pagada en el período de estudio, la fecha de salida, según corresponda, puede ser fallecimiento o censura en la fecha final del período de estudio. La muestra excluye a los asegurados que efectuaron traspaso al SPP.

No se consideraron como registros válidos a aquellas personas cuya fecha de fin de observación es inferior a enero de 2013, también se excluyen a aquellos registros cuya fecha de inicio es superior al final del período de observación. La composición de los datos válidos se muestra en la tabla 1⁴. Los datos válidos corresponden poco más de 5 millones de registros, mayoritariamente asegurados activos (87%), mayor predominancia de hombres (59%) y una muestra de 170 mil fallecidos.

Tabla 1. Datos empleados

	Absolutos	Relativos
TOTAL	5,073,561	100 %
SEXO		
Hombres	2,987,095	59 %
Mujeres	2,086,466	41 %
TIPO DE ASEGURADO		
Aseg. a Pens.	218,189	4 %
Asegurados	4,423,209	87 %
Pensionistass	432,163	9 %
CONDICIÓN DE SOBREV.		
Sobreviviente	4,903,231	97 %
Fallecido	170,330	3 %

Fuente: Elaboración propia.

⁴El proceso de obtención de los datos es el siguiente: la base de datos compuesta por todos los afiliados sanos al SNP que alguna vez fueron afiliados al sistema fueron 8.35 millones de personas; retirar información inconsistente reduce la muestra a 8.31 millones; no incluir a las personas que luego se trasladaron a una AFP desgasta la muestra a 5.36 millones; finalmente, incluir solo a aquellas personas que estuvieron expuestas al riesgo de fallecimiento en el rango de enero de 2013 a noviembre reduce la muestra al valor definitivo de 5.07 millones.

3. Resultados

3.1. Primera etapa: Estimación de la tasa bruta de mortalidad

Existen varias estrategias para la estimación de la tasa bruta de mortalidad (Ver por ejemplo London (1988), Agoston (2001) y Debon (2003)). Se consideran cuatro estimadores: un estimador basado en la propuesta de Kaplan y Meier (KM), el estimador de máxima verosimilitud (MV), el estimador provisto por el método de los momentos (MM) y el estimador actuarial (MA).

En la figura 1 se presenta la estimación de las tasas brutas de mortalidad con las cuatro definiciones antes expuestas como ratio respecto al estimador MV. Algunas reflexiones producto de la comparación se listan a continuación:

1. A excepción del método de los momentos, todas las estimaciones son similares⁵.
2. Alta dispersión en las cola inferiores, menos de 20 años y ub decrecimiento en la tasa a partir de los 95. En ambos casos este comportamiento se atribuye a la poca cantidad de datos que se tienen para la estimación en ambas colas de la distribución.
3. Así, el rango relevante se encuentra entre los 20 y 95 años, donde además se aprecia que la probabilidad de fallecimiento estimado bajo KM es consistentemente inferior que el estimador actuarial y el obtenido por máxima verosimilitud, en promedio en 1 %, tanto para hombres como mujeres.

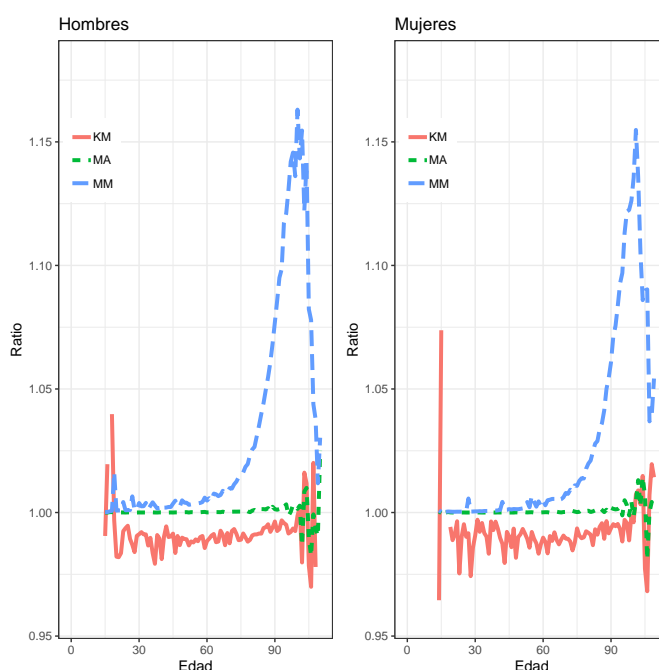


Figura 1. Tasa bruta de mortalidad como ratio respecto al estimador MV

Fuente: Elaboración propia.

Nota:

KM=Kaplan Meier; MA=Método actuarial; MM=Método de momentos; MV=Máxima verosimilitud

⁵Se puede probar que existe una relación entre el estimador MV y el de MM dado por $q_{mv} = 1 - e^{-q_{mm}}$. Matemáticamente se tiene que cuando q_{mm} tiende a cero este valor converge a q_{mv} , esto es lo que precisamente no pasa en edades avanzadas, pues la probabilidad de fallecimiento aumenta con la edad, lo cual explicaría la divergencia entre ambos métodos a medida que la edad aumenta.

En esta investigación se ha elegido el estimador KM, debido a que es el único estimador que goza de la propiedad estadística de consistencia (Kaplan Meier, 1958 y Charpentier, 2015). El principio que subyace al estimador, como se señala en London (1998), es que para que una persona sobreviva durante un año, también tendría que haber vivido en los subintervalos que se definen al interior de dicho año. Así, el estimador de la probabilidad de supervivencia es la multiplicación de las probabilidades de sobrevivir en cada subintervalo, es decir la probabilidad conjunta. Se define así el estimador producto límite siguiente:

$$\hat{q}_x = 1 - \prod_j^m (1 - \hat{q}_j)$$

donde m son los subintervalos en el intervalo $< x, x + 1]$ definidos por la ocurrencia de al menos un fallecimiento y \hat{q}_j es la probabilidad de fallecimiento en el subintervalo j . A su vez, $\hat{q}_j = d_j/n_j$; donde d_j es la cantidad de muertes observadas y n_j es el tamaño de muestra para el subintervalo j .

3.2. Segunda etapa: Graduación de la tasa bruta de mortalidad

Si el estimador obtenido en la primera etapa goza de buenas propiedades estadísticas ¿cómo es que se requiere una segunda etapa para estimar sobre lo ya estimado?, como bien señala Tomas (2012) si la probabilidad estimada en la edad i (\hat{q}_i) no influye en lo observado en j (\hat{q}_j), esto es, son independientes, entonces las tasas brutas deberían ser los estimadores finales de las verdaderas tasas. No obstante, lo razonable mas bien es pensar que las tasas serían cercanamente relacionadas a sus vecinos. Así, si la tasa bruta de mortalidad se modela como:

$$\hat{q}_i = f(x_i) + \mu_i \quad (1)$$

con μ_i representando errores aleatorios en las observaciones o variabilidad de fuentes no incluidas en x_i ; entonces la estimación \hat{q}_i cercana a \hat{q}_j debería contener información de f en x . Benjamin y Pollard (1980) recurren a la frase ‘*Natura non agit per saltum*’ para explicar el hecho que ‘las fuerzas naturales operan gradualmente y sus efectos se ponen de manifiesto de manera continua y no en saltos repentinos’⁶. De esto se sigue que las edades que se encuentran alrededor de la edad x_i pueden ser usados para aumentar la información que se tiene en esta edad y así una estimación mejorada de \hat{q}_i puede ser obtenida suavizando las estimaciones individuales.

De este modo, el siguiente paso es graduar las tasas brutas con el fin de remover alguna fluctuación aleatoria. Los enfoques que se pueden adoptar para la graduación pueden ser clasificados en paramétricos (f asume una forma predeterminada acorde a alguna ley biométrica de mortalidad) o no paramétricos (f no hace referencia a ninguna forma específica).

Debido al comportamiento particular de la tasa de mortalidad es difícil encontrar una ley de mortalidad que se ajuste razonablemente a los datos, aunque han existido esfuerzos como los realizados por Heligman y Pollard (1980) en tratar de presentar una función que modele el particular comportamiento que presenta las probabilidades de fallecimiento.

El enfoque no paramétrico, de otro lado, tiene la flexibilidad de modelar el comportamiento sin asumir un comportamiento predeterminado de la curva. Entre los enfoques más conocidos se puede señalar lo propuesto por Whittaker (1923) y Henderson (1924)⁷ y técnicas de suavizamiento de datos como ‘Spline Cúbico’ y ‘Regresión Polinómica Local’. El método de Whittaker-Henderson es el que goza de mayor popularidad en la ciencia actuarial, no obstante, su aplicación exige una elección subjetiva de algunos parámetros (Li y Chan, 2004). En este

⁶Citado en Tomas (2012), página 11.

⁷De acuerdo a esta propuesta se minimiza una función objetivo que combina los criterios de ajuste y suavidad.

trabajo se optó por la regresión polinómica local, método que se debe a Cleveland (1979) y en el que se pueden determinar criterios objetivos para la elección de los parámetros que determinan el suavizamiento.

Si en (1) el término de error se asume independiente e idénticamente distribuido con media nula y varianza finita, y $f(x_i)$ se aproxima con una función lineal alrededor de el punto x_0 , esto es en la vecindad definida por $V(x_0)$. Si además se estiman los parámetros de tal forma que las observaciones que se encuentren más cercanas a x_0 sean más importantes para la estimación, entonces el problema de optimización puede ser planteado mediante el siguiente problema de mínimos cuadrados ponderados:

$$\min_{a,b} \sum_i [y_i - a(x_0) - b(x_0)x_i]^2 K\left(\frac{x_i - x_0}{h}\right) \quad (2)$$

donde K es una función ponderadora (por ejemplo una distribución normal alrededor de x_0), h es el distancia más grande entre x_0 y algún otro punto de la vecindad $V(x_0)$. El valor estimado obtenido en esta ventana de $\hat{f}(x_0) = \hat{a}(x_0) + \hat{b}(x_0) \cdot x_0$ es el primer valor graduado, los otros se obtienen al mover la ventana alrededor de los otros valores de x . Si se reemplaza la función lineal en (2) por un polinomio de grado p , el procedimiento se generaliza entonces a lo que se conoce como una regresión polinómica local.

De acuerdo a este procedimiento se necesitan tomar tres decisiones: la forma de la función ponderadora, el grado del polinomio y el ancho de banda o tamaño de ventana en el que se realizarán las regresiones locales. En el caso del grado del polinomio, en Hastie et al. (2008) se señala que el ajuste local lineal tiende a ser sesgado en regiones de curvatura de la verdadera función, mientras que el ajuste cuadrático local generalmente es capaz de corregir este sesgo.

De otro lado, como indica Yatchew (2003), generalmente la selección de la función es menos importante que el ancho de banda. En este estudio se empleó un Kernel Tricúbico, el mismo que le da más importancia a la proximidad respecto a otras opciones igual de populares como lo son el Kernel Epanechnikov y el Kernel Gausiano⁸.

En el caso del ancho de banda, este puede ser elegido óptimamente de tal forma que el error promedio de predicción sea el menor posible. Como se sugiere en Tomas (2012) el método de validación cruzada puede ser empleado para dicho fin. De acuerdo a esta estrategia cada una de las observaciones se omite de la base de datos y es predicha con las restantes $n - 1$ observaciones. El ancho de banda se determina como aquel que retorna el menor error cuadrático medio. Se emplea la aproximación GCV⁹ como un conveniente *proxy* de la validación cruzada propiamente dicha (Hastie et al. 2008, 244).

Los anchos de banda son elegidos automáticamente de acuerdo al criterio de validación cruzada y en el intervalo de edad que va de los 20 hasta los 95 años de edad. En la figura 2 se presenta el resultado para un polinomio de grado 2 junto con las tasas brutas de mortalidad. La elección del grado de polinomio se justifica debido a que no habiendo diferencias significativas respecto a los polinominios de grado 0 y 1, se elige el polinomio de orden 2 en virtud de la recomendación de Hastie et al. (2008).

⁸El Kernel tricúbico se define como $K(\mu) = (1 - |\mu|^3)^3$ si $\mu < 1$ y 0 en otro caso

⁹Acrónimo en inglés de *Generalized Cross-Validation*.

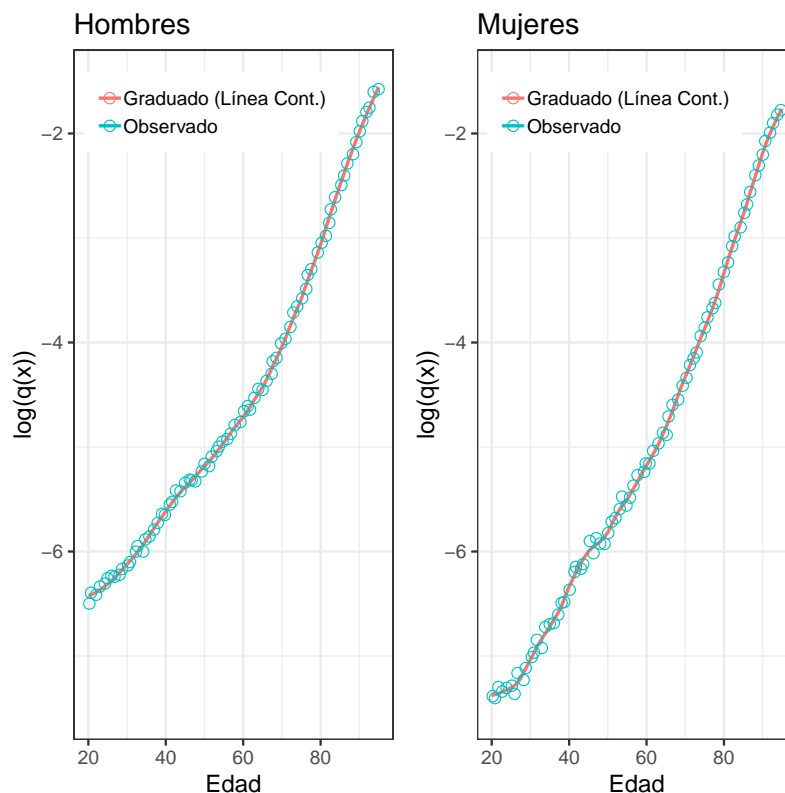


Figura 2. Graduación según regresión polinómica local
Fuente: Elaboración propia.

Culminada la graduación es válido preguntarse si en términos estadísticos esta goza de buenas propiedades. En línea con lo señalado por Forfar et al. (1988), la tradición actuarial recomienda una serie de pruebas estadísticas, aunque la aplicación de algunas de estas sean cuestionables a modo de ver de Konstantopoulos (2006). Lo que definitivamente se debería esperar de los valores graduados es que reflejen la tendencia de las tasas brutas, que la distribución de los errores esté alrededor de cero, y que los valores graduados no muestren sesgos sistemáticos como por ejemplo supere a la tasa bruta en la primera parte de la distribución y sea inferior a esta en la otra parte, o viceversa.

En la figura 3 y tabla 2 la evidencia da cuenta que la graduación efectuada sería la apropiada. La figura 3 se presentan los intervalos de confianza de las tasas brutas y los valores graduados. En la mayoría de los casos los valores graduados se encuentran al interior de los intervalos, también, al menos gráficamente, no hay evidencia que indique algún tipo de sesgo sistemático. De otro lado, de acuerdo a las pruebas presentadas en el tabla 2 se puede concluir que las tasas brutas como las graduadas se encuentran correlacionadas, que los errores en promedio se encuentran alrededor de cero y así mismo que no se puede rechazar la hipótesis nula de igualdad de medianas a los niveles usuales de significancia estadística.

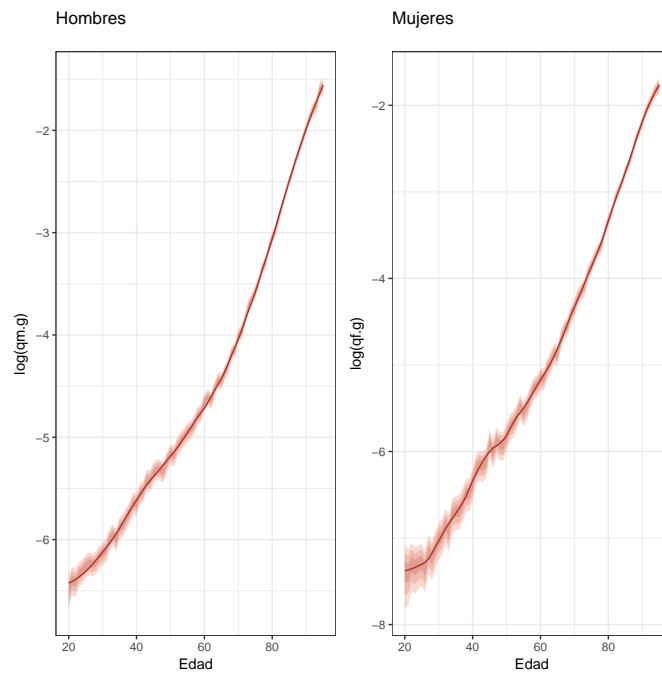


Figura 3. Tasas de mortalidad graduadas e intervalos de confianza

Fuente: Elaboración propia.

Nota:

Intervalos con una, dos y tres desviaciones estándar. El valor central corresponde a la variable graduada y los intervalos de confianza corresponde a las tasas brutas con errores estándar correspondientes al estimador Kaplan-Meier.

Tabla 2. Graduación, pruebas estadísticas. P-Values

Test	Hipótesis nula	Hombre	Mujer
Correlación	Correlación=0	0.000	0.000
Test de medias	Promedios iguales	0.998	0.997
Test de signos ^a	Medianas iguales	0.909	0.731
Test de rangos ^b	Medianas iguales	0.454	0.271

Fuente: Elaboración propia.

^a Si los signos de las diferencias no nulas son similares la mitad de estas son aproximadamente positivas y la otra mitad negativas.

^b Corresponde a la prueba de signos de Wilcoxon, además del signo el test también considera el tamaño de la diferencia (rango).

3.3. Tercera etapa: Estimación en las colas

En el caso de la cola inferior se empleó como referencia la tabla de mortalidad estimada por el INEI para el período 2010-2015, INEI (2009). La estrategia en este tramo consta de tres etapas. En la primera se estima las probabilidades anuales de fallecimiento a partir de la información reportada en la tabla abreviada del INEI¹⁰. La estrategia en este tramo constó de los siguientes pasos:

1. Las probabilidades quinquenales se anualizaron con la siguiente expresión: $q_x = 1 - (1 -_5q_x)^{1/5}$.
2. El valor anualizado se escala de tal forma que coincida con la tasa graduada conocida a los 20 años de edad.
3. Los valores restantes se completan mediante una interpolación que emplea una regresión polinómica local entre el logaritmo de la probabilidad y la edad para cada sexo.

Vale la pena señalar que las personas menores de 20 años representa un porcentaje pequeño respecto a la población, representa alrededor del 1% tanto de pensionistas como afiliados¹¹.

De otro lado, en el caso de la cola superior, se extrapolaron las tasas empleando para ello alguna función biométrica¹². Se ensayaron cinco funciones: Exponencial, Gompertz, Makeham, Heligman-Pollard y Kannisto. Los parámetros de las mismas se estimaron usando la información de las probabilidades graduadas bajo la estrategia de mínimos cuadrados no lineales.¹³

Los modelos elegidos fueron aquellos que reportaron los menores errores de estimación empleando la menor cantidad de parámetros (criterio de parsimonia). Tanto en hombres como mujeres se eligió la función conocida como Kannisto. Para la extrapolación se le dio más importancia a la muestra más cercana al horizonte de predicción. En los anexos se muestran los indicadores de Akaike y Bayesiano que permitieron elegir los modelos y los parámetros estimados de los modelos elegidos.

Una vez determinado el modelo con el cual se van a extrapolar las probabilidades, queda por definir hasta que edad se realizará dicha extrapolación. En otras palabras, se necesita definir el último valor de la tabla. Como se muestra a continuación la elección de dicho valor respecto a otro tendría efectos marginales en la esperanza de vida, y por tanto en las estimaciones actuariales.

En la tabla 3 se presentan las esperanzas de vida a los 20 y 65 años asumiendo tres diferentes topes máximos de edad: 110, 115 y 120 años. No existen diferencias significativas entre los distintos escenarios. En el caso de los hombres, la esperanza de vida es de 77.9 si se mide a los 20 años y de 84.7 si se mide a los 65 años; en el caso de las mujeres la misma estimación es de 83.0 en el primer caso y de 86.9, respectivamente. Las diferencias en expectativa de vida entre los distintos escenarios se notan recién a partir del tercer o cuarto decimal.

Siendo así, se propone fijar el infinito actuarial en 110 años y recomendar, que de darse el caso de contar con una persona con una edad superior a los 110 años, se la trate como si tuviera esta edad. Esta última recomendación es consistente con tener una probabilidad de fallecimiento menor a 1 en este umbral. En la figura 4 se muestran los resultados de la extrapolación hasta los 110 años.

¹⁰En una tabla abreviada las probabilidades se presentan por intervalos de edades a diferencia de una tabla completa en la que las probabilidades son por cada año.

¹¹Valores estimados para el año 2017 en el caso del régimen del Decreto Ley 19990.

¹²El comportamiento particular de las probabilidades en edades tempranas y la muestra reducida en este rango no ha permitido emplear la estrategia de estimación de una función biométrica en la cola inferior.

¹³Estimaciones realizadas con el algoritmo Levenberg-Marquardt.

Tabla 3. Esperanza de vida

A los 20 años ^{a,b}		
Edad máxima ^c	Hombre	Mujer
$\omega=110$	77.8893140935905	82.9633197233986
$\omega=115$	77.8894596902662	82.9639907563457
$\omega=120$	77.8894629075012	82.9640094519793
A los 65 años ^{a,b}		
Edad máxima ^c	Hombre	Mujer
$\omega=110$	84.6733577409578	86.8560238265777
$\omega=115$	84.6735376224666	86.8567771715841
$\omega=120$	84.6735415972901	86.8567981605122

Fuente: Elaboración propia.

^a La esperanza de vida a la edad x se define como $x + \sum_{i=x+1}^{\omega} l_i/l_x + 0,5$, donde l_x es el número estimado de sobrevivientes a la edad x .

^b Se emplea la función Kannisto para extrapolar hasta el infinito actuarial indicado.

^c Infinito actuarial, valor a partir del cual se considera no hay sobrevivientes.

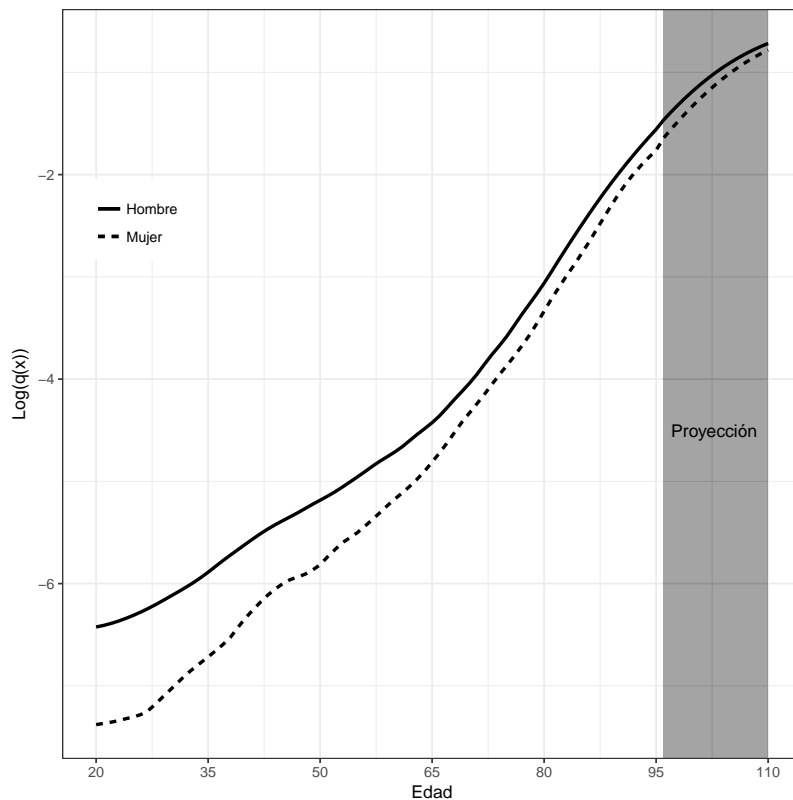


Figura 4. Estimación en cola superior

Fuente: Elaboración propia

3.4. Análisis comparativo

En la figuras 5 y tabla 4 se presentan los resultados de la nueva tabla y se compara respecto a la tabla vigente (SNP 2005) y también con la tabla recientemente estimada y aprobada por la Superintendencia de Banca y Seguros para los afiliados al Sistema Privado de Pensiones (SPP). Las probabilidades q_x siguen el comportamiento típico esperado, tasas altas en edades tempranas y crecimiento a tasas decrecientes en edades avanzadas. Esto último es consistente con los resultados encontrados por otras investigaciones (Thatcher, 1999 y Horiuchi-Wilmoth, 1998, por ejemplo). Por su parte, la expectativa de vida de los hombres y mujeres medido a los 65 años en la nueva tabla es de 84.7 y 86.9, respectivamente; significando un aumento y una reducción de 2.1 y 2.6 años respecto a la tabla vigente; la brecha entre sexos es consistente con lo observado en el sistema privado.

La comparación entre la tabla SNP 2017 y la SPP 2017 permite verificar que las probabilidades de fallecimiento son mayores en el caso de los afiliados al SNP. Este resultado es esperado teniendo en cuenta que los afiliados al SNP tienen menores ingresos que sus pares del SPP y sabiendo que la evidencia empírica internacional señala que las personas de menores niveles socioeconómicos tienen menor expectativa de vida que sus pares (Ver por ejemplo Duleep (1986) y Deaton y Paxson (1998)).

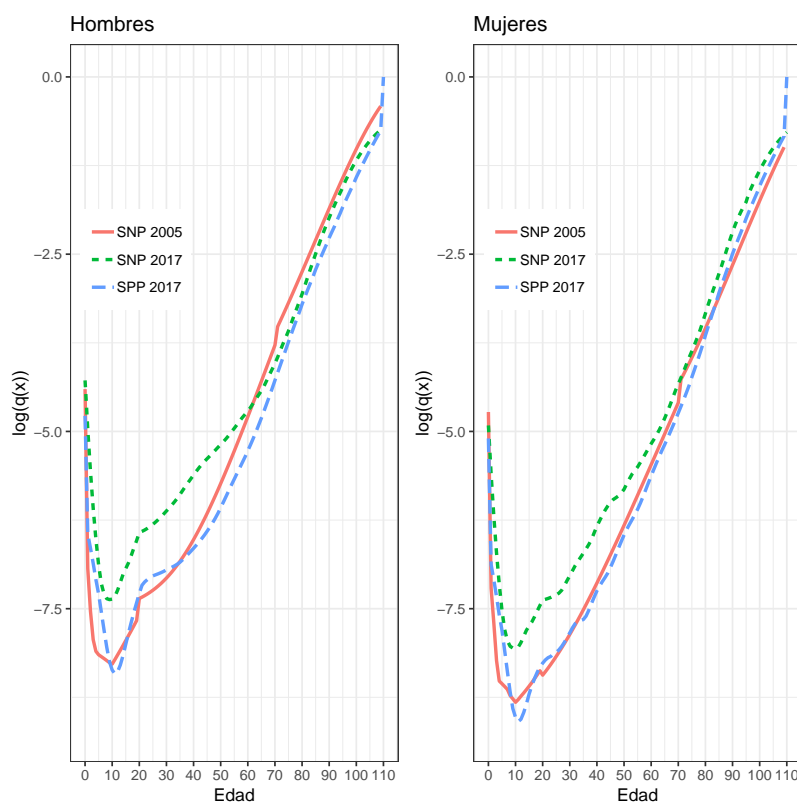


Figura 5. Probabilidades de fallecimiento según tablas (en logaritmos)

Fuente: Elaboración propia

Tabla 4. Esperanza de vida según edad y tabla

Edad	Hombres			Mujeres		
	SNP17 ¹	SP2005 ²	SPP17 ³	SNP17 ¹	SP2005 ²	SPP17 ³
0	74.9	77.3	81.7	81.3	85.7	86.3
20	77.9	78.7	83.1	83.0	86.8	87.3
65	84.7	82.6	86.5	86.9	89.3	89.5

Fuente: Elaboración propia.

¹ Propuesta de tabla para personas sanas.

² Tabla de mortalidad actual para sanos.

³ Tabla de mortalidad para personas sanas empleada en el sistema privado de pensiones.

Respecto a las implicancias financieras, una menor expectativa de vida en el caso de las mujeres, implicaría una menor previsión actuarial tanto por la titularidad como por la condición de viudez. De otro lado, un incremento en la esperanza de vida del hombre tiene un efecto contrario pero solo en el caso de la titularidad, pues en el caso del régimen pensionario del D.L. 19990, solo las potenciales viudas podrían ser beneficiarios de una pensión, más no los potenciales viudos.

En la tabla 5 se presenta el efecto sobre las reservas actuariales debido al cambio de tabla de mortalidad. Los cálculos corresponden al ejercicio actuarial 2017 y de estos se desprende que la reserva neta disminuiría en $S/2,032$ millones de soles, esto es, una reducción de 1,8%. Esta reducción neta se explica básicamente por el efecto en las reservas en el caso de las mujeres, pues este se reduce en $S/4,217$ millones de soles mientras que la reserva de los hombres aumenta en $S/2,185$ millones.

Tabla 5. Efecto sobre las reservas actuariales 2017 (en millones de soles)

	Efecto total			
	SP2005	SNP17	Var. en Niv.	Var. Porc.
Hombres	73,096	75,281	2,185	3.0%
Mujeres	40,330	36,113	-4,217	-10.5%
Total	113,426	111,394	-2,032	-1.8%

Fuente: Elaboración propia.

3.5. Robustez de los resultados

En esta sección se presenta la sensibilidad de los resultados bajo tres escenarios: (i) Se evalúa el impacto de usar la tasa bruta de mortalidad empleando el criterio actuarial en lugar del estimador no paramétrico de Kaplan y Meier; (ii) el efecto de usar la técnica de graduación conocida como Whittaker-Henderson en lugar de la regresión polinómica local; y (iii) las implicancias de usar la misma metodología empleada en la tabla SP2005 (Detalles metodológicos de esta última pueden ser consultadas en Melinzky, 2006). En los dos primeros casos se trata de un efecto que sólo puede ser atribuido a dicha modificación, pues todos los demás pasos son los mismos planteados en la presente propuesta. En el caso del tercer ejercicio, este debe ser leído como un efecto combinado, pues la metodología SP2005 contempla el uso de el método actuarial para

estimar la tasa bruta, la graduación según Wittaker-Henderson y la primera ley de Makeham para la extrapolación en edades avanzadas, entre las características más saltantes.

En la figura 6 se presentan las expectativas de vida como ratio respecto a la nueva tabla de vida propuesta, la SNP2017. En general se tiene que las diferencias aunque pequeñas, son distintas dependiendo del tipo de estimación. De esta manera se tiene que la metodología SP2005 es la que más discrepa de la SNP2017 en términos relativos, siendo las brechas más importantes en las edades más avanzadas, así, a la edad de 105 años mientras la expectativa de vida se estima en 106.8 y 107.0 para hombres y mujeres, respectivamente, empleando la estrategia empleada en la SP2005 se tienen esperanzas de 106.3 y 106.5.

De otro lado, en el caso de la tasa bruta actuarial, también se tienen menores expectativas de vida, en este caso las diferencias se acortan hasta converger en edades avanzadas por la que la diferencia más amplia se encuentra a los 20 años, donde, mientras la expectativa de vida se estima en 77.8 y 82.8 para hombres y mujeres, respectivamente, los mismos indicadores en la SP2017 son de 78.0 y 83.0. Finalmente, en el caso de la técnica de graduación, prácticamente no hay diferencia entre usar el método de Whittaker-Henderson y la regresión polinómica local propuesta en este documento. Tal vez, la ventaja de este último sobre el primero es que se puede implementar de manera automática sin necesidad de elegir arbitrariamente los parámetros que necesita el primer método para su estimación.

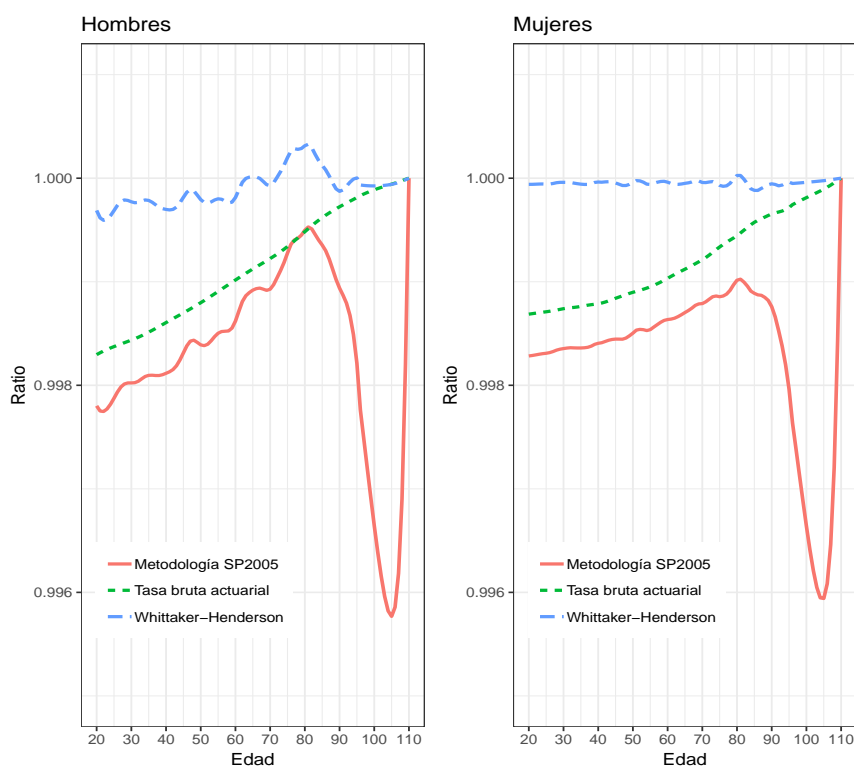


Figura 6. Esperanza de vida como ratios respecto a SNP2017

Fuente: Elaboración propia.

4. Conclusiones

En este estudio se emplearon poco más de 5 millones de registros para la estimación de tablas de vida estática para la población sana e inválida afiliada al Sistema Nacional de Pensiones. La información proviene de los registros administrativos de la Oficina de Normalización Previsional acumulados en los últimos 5 años, la misma que se complementa con información proveniente del Registro Nacional de Identificación y Estado Civil.

La estrategia de estimación consistió en estimar primero la tasa bruta de mortalidad empleando el estimador producto límite sugerido por Kaplan-Meier. Luego, la graduación o suavizamiento se consiguió con la técnica de regresión polinómica local. Finalmente, para lidiar con la insuficiencia de datos en las colas, las tasas en estos tramos fue calculado con una estrategia diferenciada: en la cola inferior, edades inferiores a los 20, se ajustó la estimación realizada por el INEI para toda la población; en el caso de la cola superior, edades superiores a los 95, se extrapoló hasta los 110 usando la función biométrica de Kannisto tanto para hombres como mujeres. Se encontró que incluso variando el tope máximo a 115 o 120, el indicador de expectativa de vida no cambia significativamente.

Como resultado de la estimación se encontró que, respecto a la tabla SP2005, las probabilidades de fallecimiento en el caso de los hombres disminuyen y las de las mujeres aumentan. Como consecuencia de ello la brecha en esperanza de vida entre mujeres y hombres se reduce de 7 a 2 años cuando este indicador mide a los 65. La implicancia directa de este resultado es que el dinero necesario para financiar las pensiones de las mujeres se reduciría y la de los hombres aumentaría. En el año 2017 se estima que la deuda actuarial neta se reduciría en S/2 mil millones de soles, esto es, en un 2% respecto a lo estimado oficialmente para dicho año.

Ampliaciones de este estudio también deberían explorar estimaciones dinámicas de la tabla. Esto es, tener en cuenta que las probabilidades de sobrevivencia o fallecimiento podrían estar cambiando entre distintas generaciones. El tratamiento estándar que se le suele dar a esta exigencia típicamente sigue lo planteado por Lee y Carter (1992), sin embargo, este o en general cualquier método que requiera como insumo de tablas históricas, no se pueden realizar en este caso, pues no se cuenta con ninguna historia y, las que se podrían estimar usando datos propios se encuentran lejos del mínimo de 30 puntos que suele exigir la modelación de series de tiempo. Las tablas aquí presentadas son las primeras tablas estimadas con la población de afiliados al SNP, por lo que cualquier esfuerzo de ajuste en las tablas basada en la historia, debe contemplar esta limitación y señalar los posibles sesgos que se podrían estar cometiendo al usar alguna estrategia retrospectiva.

Referencias bibliográficas

- [1] Bueno, J., Escuder, R., Pavia, J. y Guillén, M. (2008), *Determinación de los Tantos Brutos de Mortalidad*, Anales del Instituto de Actuarios Españoles, N° 14 3ra época, 109-133.
- [2] Charpentier, A. (Ed.) (2014). *Computational Actuarial Science With R*, CRC Press.
- [3] Cleveland, W.S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 74(368), 829-836.
- [4] Deaton, A. y Paxson, C. (1998). Aging and Inequality in Income and Health. *The American Economic Review*, 88(2), 248-253.
- [5] Duleep, H. O. (1986), Measuring the Effect of Income on Adult Mortality Using Logitudinal Administrative Record Data. *Journal of Human Resources*, 21(2), 238-251.

- [6] Forfar, D.O., McCutcheon, J.J. y Wilkie, A.D. (1988). On Graduation by Mathematical Formula. *Journal of the Institute of Actuaries*, 115(1), 1-149.
- [7] Gil Fana, J., Heras Martínez, A. y Vilar Zanón, J. L. (s.f). *Matemática de los Seguros de Vida*. Fundación MAPFRE Estudios.
- [8] Hastie, T., Tibshirani, R. y Friedman, J. (2008). *The Elements of Statistical Learning*. Second Edition. New York, USA: Springer.
- [9] Heligman, L. y Pollard, J. (1980). The Age Pattern of Mortality. *Journal of the Institute of Actuaries*, 107(1), 49-80.
- [10] Horiuchi, S. y Wilmoth, J.(1998). Deceleration in the Age Pattern of Mortality at Older Ages. *Demography*, 35(4), 391-412.
- [11] I. Massons, J. M. D. (1992). Una aplicación del análisis de la supervivencia en ciencias de la salud. *Anuario de psicología/The UB Journal of psychology* (55), 109-142.
- [12] INEI (2009). *Perú: Estimaciones y Proyecciones de Población 1950-2050*. Boletín de Análisis Demográfico 36.
- [13] Konstantopoulos, T. (2006). *Notes on Survival Models*, School of Mathematical and Computer Sciences Heriot-Watt University,
- [14] Lee, R. y Carter, L. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87(419), 659-671.
- [15] Li, Siu-Hang y Chan, Wai-Sum (2004). Estimation of Complete Period Life Tables for Singaporeans. *Journal of Actuarial Practice*, 1(11) 129-146.
- [16] London, D. (1988). *Survival Models and Their Estimation*. Second Edition. USA: ACTEX Publications.
- [17] Melinsky, E. (2006) *Elaboración de Tablas de Mortalidad en Perú*. Lima, Perú: manuscrito.
- [18] Sánchez, J. (2002). Estimación Actuarial Versus Estimación por el Método de los Momentos para la Probabilidad de Muerte. *Anales del Instituto de Actuarios Españoles*, 8, 111-135.
- [19] Thatcher, A. (1999), The Long-Term Pattern of Adult Mortality and the Highest Attained Age. *Journal of the Royal Statistical Society*, 162(1), 5-43.
- [20] Tomas, J. (2012). Univariate Graduation of Mortality by Local Polynomial Regression. *Bulletin Francais d'Actuariat* 12, (23), 5-58.
- [21] Yatchew, A.(2003), *Semiparametric Regression for the Applied Econometrician* United States of America: Cambridge University Press.

Anexos

Tabla 6: Elección de modelo para extrapolación en cola superior.

Hombres - AIC					
Edades	Gompertz	Makeham	H-P	Kannisto	Exponencial
30-95	-600.16	-650.06	-596.71	-720.02	-619.26
40-95	-506.96	-542.11	-504.37	-625.60	-515.16
50-95	-421.79	-435.99	-421.49	-508.71	-415.14
60-95	-339.30	-337.86	-347.18	-389.78	-323.12
Mujeres - AIC					
Edades	Gompertz	Makeham	H-P	Kannisto	Exponencial
30-95	-635.96	-647.54	-641.18	-685.63	-628.45
40-95	-531.61	-539.34	-536.51	-576.62	-522.86
50-95	-431.58	-432.99	-437.13	-464.15	-420.21
60-95	-332.43	-330.44	-339.63	-352.83	-321.71

Tabla 7: Elección de modelo para extrapolación en cola superior.

Hombres - BIC					
Edades	Gompertz	Makeham	H-P	Kannisto	Exponencial
30-95	-593.59	-641.31	-590.14	-711.26	-610.50
40-95	-500.88	-534.01	-498.29	-617.50	-507.06
50-95	-416.30	-428.68	-416.01	-501.39	-407.82
60-95	-334.55	-331.52	-342.43	-383.44	-316.79
Mujeres - BIC					
Edades	Gompertz	Makeham	H-P	Kannisto	Exponencial
30-95	-629.39	-638.79	-634.61	-676.87	-619.69
40-95	-525.53	-531.24	-530.43	-568.52	-514.76
50-95	-426.10	-425.68	-431.64	-456.84	-412.89
60-95	-327.68	-324.10	-334.88	-346.49	-315.38

Tabla 8: Modelos elegidos para extrapolación

Parámetros	Hombre ^a	Mujer ^a
γ_1	2.213e-06*** (4.875e-08)	1.569e-06*** (6.348e-08)
γ_2	1.246e-01*** (2.135e-04)	1.256e-01*** (3.892e-04)
γ_3	4.352e-03*** (2.396e-04)	2.377e-03*** (3.831e-04)
$\hat{\sigma}_\epsilon$	0.0008276	0.001386
Grados de Libertad	48	48

^a Corresponde a la función biométrica de Kannisto:

$$\hat{q}_x = 1 - \exp\left(-\left(\frac{\gamma_1 \cdot \exp(\gamma_2 \cdot x)}{1 + \gamma_1 \cdot \exp(\gamma_2 \cdot x)} + \gamma_3\right)\right).$$

**** significancia al 0.0001

Tabla 9: Tabla de vida Hombres. Población Sana. (SNP2017)

Edad	q_x	Edad	q_x	Edad	q_x
0	0.0138	37	0.0031	74	0.0254
1	0.0070	38	0.0033	75	0.0278
2	0.0038	39	0.0035	76	0.0309
3	0.0023	40	0.0036	77	0.0344
4	0.0015	41	0.0038	78	0.0381
5	0.0010	42	0.0040	79	0.0421
6	0.0008	43	0.0042	80	0.0468
7	0.0007	44	0.0044	81	0.0525
8	0.0006	45	0.0046	82	0.0590
9	0.0006	46	0.0048	83	0.0661
10	0.0006	47	0.0050	84	0.0739
11	0.0007	48	0.0052	85	0.0825
12	0.0007	49	0.0054	86	0.0919
13	0.0008	50	0.0056	87	0.1021
14	0.0009	51	0.0058	88	0.1131
15	0.0010	52	0.0061	89	0.1248
16	0.0011	53	0.0064	90	0.1372
17	0.0012	54	0.0067	91	0.1504
18	0.0013	55	0.0070	92	0.1643
19	0.0014	56	0.0074	93	0.1790
20	0.0016	57	0.0078	94	0.1945
21	0.0016	58	0.0082	95	0.2106
22	0.0017	59	0.0086	96	0.2309
23	0.0017	60	0.0090	97	0.2495
24	0.0018	61	0.0095	98	0.2687
25	0.0018	62	0.0100	99	0.2884
26	0.0019	63	0.0107	100	0.3082
27	0.0019	64	0.0113	101	0.3282
28	0.0020	65	0.0120	102	0.3481
29	0.0021	66	0.0128	103	0.3678
30	0.0022	67	0.0138	104	0.3871
31	0.0023	68	0.0150	105	0.4059
32	0.0024	69	0.0162	106	0.4240
33	0.0025	70	0.0176	107	0.4414
34	0.0026	71	0.0192	108	0.4579
35	0.0028	72	0.0212	109	0.4735
36	0.0029	73	0.0233	≥ 110	0.4882

Tabla 10: Tabla de vida Mujeres. Población Sana. (SNP2017)

Edad	q_x	Edad	q_x	Edad	q_x
0	0.0073	37	0.0014	74	0.0191
1	0.0037	38	0.0015	75	0.0209
2	0.0020	39	0.0016	76	0.0229
3	0.0012	40	0.0018	77	0.0253
4	0.0008	41	0.0019	78	0.0280
5	0.0005	42	0.0021	79	0.0315
6	0.0004	43	0.0022	80	0.0356
7	0.0004	44	0.0024	81	0.0402
8	0.0003	45	0.0025	82	0.0450
9	0.0003	46	0.0026	83	0.0503
10	0.0003	47	0.0026	84	0.0558
11	0.0003	48	0.0027	85	0.0624
12	0.0003	49	0.0028	86	0.0699
13	0.0004	50	0.0030	87	0.0789
14	0.0004	51	0.0032	88	0.0889
15	0.0004	52	0.0035	89	0.1002
16	0.0005	53	0.0037	90	0.1122
17	0.0005	54	0.0039	91	0.1245
18	0.0005	55	0.0041	92	0.1364
19	0.0006	56	0.0043	93	0.1483
20	0.0006	57	0.0046	94	0.1600
21	0.0006	58	0.0050	95	0.1715
22	0.0006	59	0.0053	96	0.1935
23	0.0006	60	0.0057	97	0.2109
24	0.0007	61	0.0060	98	0.2290
25	0.0007	62	0.0064	99	0.2479
26	0.0007	63	0.0069	100	0.2673
27	0.0007	64	0.0075	101	0.2870
28	0.0008	65	0.0081	102	0.3071
29	0.0008	66	0.0089	103	0.3273
30	0.0009	67	0.0098	104	0.3473
31	0.0009	68	0.0109	105	0.3672
32	0.0010	69	0.0120	106	0.3867
33	0.0011	70	0.0132	107	0.4056
34	0.0011	71	0.0144	108	0.4239
35	0.0012	72	0.0158	109	0.4414
36	0.0013	73	0.0174	≥ 110	0.4580