

Libro III, Título X Tablas de Mortalidad, Anexos**Anexo N° 4 Nota técnica tablas de mortalidad
B-2006 (Hombres y mujeres) Y MI-2006
(Hombres y mujeres)****I. Introducción**

Conforme a lo dispuesto en los artículos 55 y 65 del D.L. N° 3.500, de 1980 y en el artículo 20 del D.F.L. N°251, de 1931, las Superintendencias de Valores y Seguros y de Pensiones, desarrollaron las tablas de mortalidad, MI-2006 (hombres y mujeres) para pensionados por invalidez y beneficiarios inválidos de pensión de sobrevivencia, y B-2006 (hombres y mujeres) para beneficiarios no inválidos de pensión de sobrevivencia.

Las tablas MI-2006 y B-2006 deberán ser utilizadas para el cálculo de los retiros programados, del aporte adicional y de las reservas técnicas por parte de aseguradoras del segundo grupo, que mantengan obligaciones por la contratación de seguros de renta vitalicias del D.L. N°3.500, de 1980.

El proceso de construcción de las tablas antes señaladas se dividió en 4 etapas:

- Obtención y depuración de datos
- Cálculo de expuestos y determinación tasas brutas de mortalidad
- Ajuste de tasas brutas de mortalidad y Aplicación de test estadísticos
- Cálculo de los factores de mejoramiento y márgenes de seguridad

II. Obtención y depuración de base de datos

La base de datos utilizada para la construcción de las tablas se obtuvo de tres fuentes de información que involucran al sistema previsional chileno:

- a. Base de pólizas de Renta Vitalicia SVS
- b. Base de pensionados del Sistema DL. 3.500 SP
- c. Base de pensionados Sistema IPS (ex cajas de previsión)

Con la fusión de la Base de pólizas de Renta Vitalicia SVS y Base de pensionados del Sistema DL. 3.500 SP se conformo una Base de Datos única del Sistema de Pensiones regulado por el D.L. N° 3.500

1. Base de pólizas de Renta Vitalicia SVS**- Conformación de la Base de Datos**

Se utilizaron los datos que trimestralmente envían las compañías de seguros, y reaseguros, correspondientes al Stock de pólizas y siniestros, vigentes o hayan dejado de estarlo, conforme a en la Letra A del Título II del Libro II y en la Letra D del Título I del Libro IV.

Los datos recibidos incluyen antecedentes de la póliza de renta vitalicia, o del siniestro antes mencionado, de los afiliados causantes de dichas pólizas y de sus beneficiarios.

- Pólizas incluidas

Se utilizó el stock de datos al 31 de diciembre de 2005. Se incluyeron todas las pólizas de RV inmediata y RV diferida, en estas últimas tanto que se haya iniciado el pago de la renta vitalicia, como aquellas en que aún no comienzan a devengarse las rentas. Además se consideraron los siniestros de invalidez y sobrevivencia, ocurridos con anterioridad 31/12/1987 (Letra D del Título I del Libro IV).

- Registros excluidos

No se incluyó en el análisis las pólizas que corresponden a aceptaciones de reaseguro, tanto de las reaseguradoras como las aceptaciones entre aseguradoras.

Los datos de la compañía Le Mans fueron excluidos por encontrarse en quiebra. Así mismo se excluyeron de la base de trabajo los beneficiarios designados, por tratarse de registros cuya única función dentro de la póliza es asignar la reserva por la pensión no percibida por un rentista fallecido, cuando no existen beneficiarios con derecho a pensión, siendo en muchos casos información que no corresponde a un beneficiario "natural".

- Depuración base de datos de personas duplicadas

Los registros duplicados se generan debido a la posibilidad que un causante de pensión tenga más de una póliza de renta vitalicia, con la consecuente repetición del grupo familiar en cada póliza o por el hecho que en un grupo familiar ambos cónyuges mantengan póliza de Renta Vitalicia, en cuyo caso los hijos serían beneficiarios de ambos causantes, duplicándose de esta forma estos beneficiarios. En estos casos, se estableció como regla general mantener los datos de la póliza más antigua.

- *Verificación de datos en el Registro Civil*

Con el objeto de validar la información a utilizar, los datos fueron enviados al registro civil para verificar fechas de nacimiento, fallecimiento y sexo de las personas que constituyen esta base. La base de datos enviada a verificación consistió en 605.516 registros correspondientes a causantes de renta vitalicias mujeres, causantes hombres inválidos, y beneficiarios.

2. Base de pensionados del Sistema D.L. N° 3.500 SP

- *Conformación de la Base de Datos*

La información proporcionada por las Administradoras de Fondos de Pensión para estos efectos, corresponde a afiliados declarados inválidos definitivos desde 1981 hasta 2005, potenciales beneficiarios de pensión de sobrevivencia, es decir aquellos declarados por los afiliados que se encuentran percibiendo pensión de vejez o invalidez en la modalidad de retiros programados y beneficiarios acogidos a pensión de sobrevivencia en dicha modalidad.

- *Verificación de datos en el Registro Civil*

Las Administradoras, en forma previa a la entrega de la información señalada verificaron con el Registro Civil la condición de vivos o muertos para todos aquellos beneficiarios potenciales inválidos y no inválidos, vivos o muertos, que adquirieron la calidad de beneficiarios debido a que el afiliado con el cual tienen alguna relación de parentesco se acogió a pensión entre el 01.05.1981 y el 31.12.2005, y dicho afiliado se encuentre vivo al 31.12.2005.

- *Depuración base de datos de personas duplicadas*

Al igual que con la base de datos de la SVS, se chequearon duplicados entre beneficiarios e inválidos ya que existe la posibilidad de que uno de ellos sea beneficiario más de una vez, por ejemplo si su padre y su madre están pensionados en la modalidad retiro programado.

3. Base de pensionados Sistema IPS

- *Conformación de la Base de Datos*

Se utilizó la base de datos proporcionada por el IPS conformada por todos los pensionados vivos y muertos al 31/12/2005, que al 01/01/1999 se encontraban vivos y con derecho a pensión o que entraron al sistema con ese derecho, entre las fechas señaladas. Del mismo modo, se incorporó el dato de todos aquellos pensionados que, habiendo estado vivos y vigentes al 01/01/ 1999, hayan salido del sistema entre estas fechas.

- *Registros excluidos*

No se incluyó para el análisis y posterior elaboración de las tablas, los pensionados de Accidentes del Trabajo y Enfermedades Profesionales.

- *Depuración base de datos de personas duplicadas*

La depuración de la base de datos consistió en dejar una sola vez cada afiliado causante pensionado de invalidez y cada beneficiario inválido o no inválido.

Los registros duplicados se generan por la posibilidad que un causante de pensión sea pensionado de más de una caja, o por el hecho que en un grupo familiar ambos cónyuges sean pensionados del IPS, en cuyo caso los hijos serían beneficiarios de ambos causantes, duplicándose de esta forma estos beneficiarios.

- *Verificación de datos en el Registro Civil*

Con el objeto de validar los datos a utilizar, se pareó con el Registro Civil una muestra representativa de 90.000 casos, obteniéndose un resultado satisfactorio.

4. Base de Datos del Sistema de Pensiones regulado por el D.L. N° 3500

- *Conformación de la Base de Datos*

Se fusionaron las bases de datos provenientes de la Superintendencia Valores y Seguros y Superintendencia de Pensiones, conformando una base representativa del Sistema de Capitalización Individual de Chile.

- *Depuración base de datos de personas duplicadas*

Al igual que en las bases de datos anteriores existe la posibilidad de datos repetidos, ya que una persona se puede traspasar de retiro programado a renta vitalicia, o mantener una pensión en ambas modalidades. Por este motivo la base se depuró en base al RUT, conformando un registro único por persona.

III. Metodología de cálculo de tasas brutas y expuestas

1. Nomenclatura y definiciones

En este informe se utiliza la notación actuarial generalmente aceptada.

- Θ_x número de muertos observados a la edad x .
- E_x cantidad de expuestos al riesgo en la edad x .
- $q_{x,t}$ probabilidad de que una persona de edad x al momento t muera antes de cumplir la edad $x+1$ al momento $t+1$.
- $q^{\circ}_{x,t}$ tasa bruta de mortalidad; valor observado de $q_{x,t}$.
- $q^{\prime}_{x,t}$ se refiere a la tasa de mortalidad ajustada
- $r_{x,t}$ factor de mejoramiento del valor $q_{x,t}$ del momento $t-1$ al momento t .

2. Cálculo de las tasas brutas de mortalidad

El cálculo se efectúa de la siguiente manera:

$$q^{\circ}_{x,t} = \Theta_x / E_x.$$

Donde Θ_x es la cantidad de muertes ocurridas y E_x es la cantidad de expuestos al riesgo en la edad x .

3. Cálculo de Expuestos al Riesgo

Primero es necesario calcular la edad asegurada "IA" que corresponde a:

IA = Fecha exacta bautizo de la póliza - fecha exacta de nacimiento.

La fecha exacta se expresa en números decimales, para posteriormente aproximarla al número entero más cercano (edad actuarial).

Luego se recalcula la fecha de nacimiento (VYB) teniendo en cuenta la nueva edad a la toma del seguro.

$$VYB = CYI - IA$$

Donde CYI es el año calendario en que se bautizó la póliza (sin mes ni días).

Una vez obtenidos los valores de IA y VYB se puede calcular el siguiente vector:

$$v_i = [y_i, z_i, \theta_i, \phi_i]$$

Donde:

	Nombre Variable	Forma de Cálculo
y_i	edad en que comienza la observación	Año en que comienza la observación - VYB
z_i	edad en que sale del periodo de observación	Año en que termina la observación - VYB
θ_i	edad exacta de muerte	IA + muerte exacta - Bautizo exacto de la póliza
ϕ_i	edad de renuncia	Año en que renuncia - VYB

Una vez tenido esos datos se pueden calcular los expuestos y fallecimientos por año.

El cálculo se realiza de la siguiente forma:

Expuestos a la edad $x = A$ la suma de todos los individuos que cumplan con estos requisitos:

$$(y_i < x + 1); (z_i > x + 1); (\theta_i = 0 \text{ ó } x \leq \theta_i); (\varphi_i = 0 \text{ ó } x \leq \varphi_i)$$

Fallecimientos a la edad x = el subconjunto de expuestos a la edad x que cumplen con

$$x < \theta_i \leq x+1$$

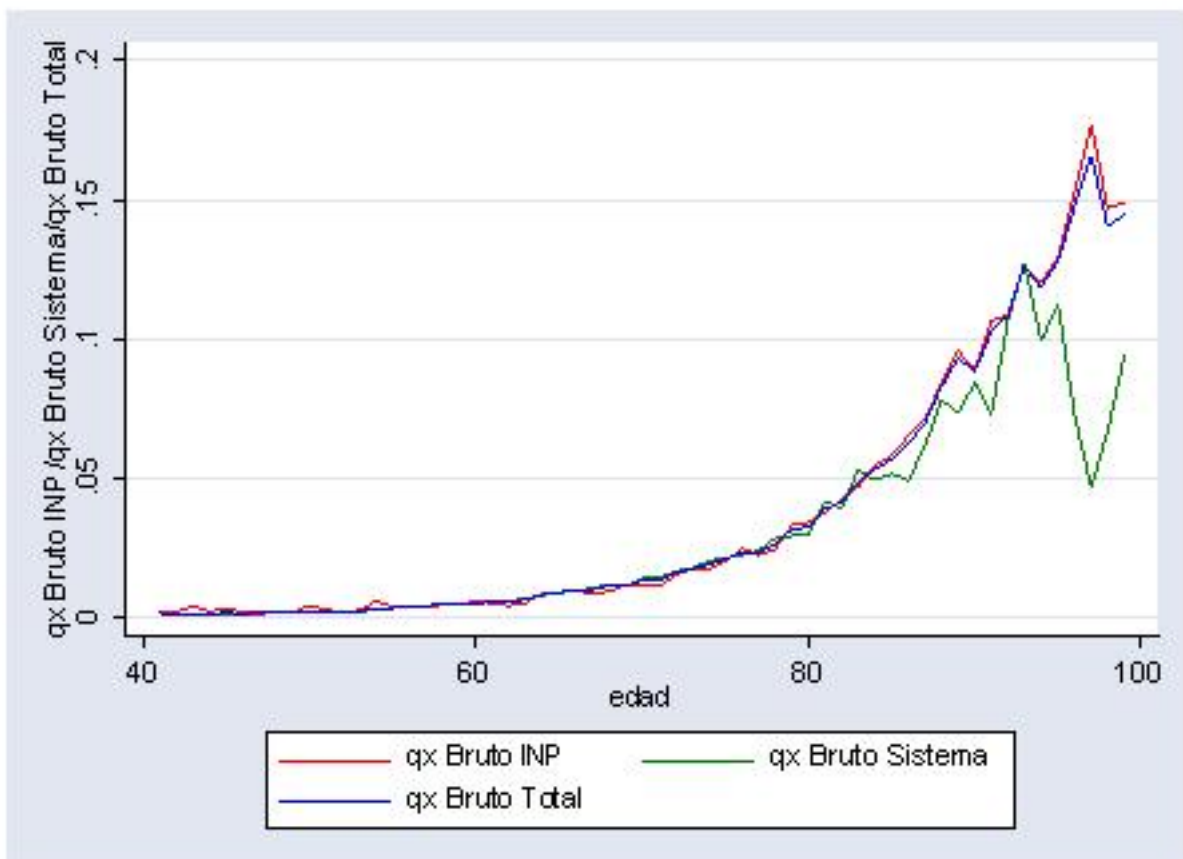
4. Cálculo de expuestos de beneficiarias

Para este cálculo se consideró un período de cuatro años, con cortes entre el 2001 y el 2004 inclusive, tanto para la información del sistema de capitalización individual como, para el sistema de reparto, IPS.

Además, para el caso del IPS se consideró dentro de ese período a las beneficiarias con una última pensión mayor a 70.000 pesos. Lo anterior debido a que la pensión mínima de un pensionado mayor a 70 años es de alrededor de 105.000 pesos y a que las cónyuges les corresponde por lo general un 60% de la pensión que recibía el imponente fallecido. Este corte es equivalente al utilizado en la construcción de la tabla RV 2004, que se efectuó para utilizar información del IPS de imponentes que hubieran podido contratar una Renta Vitalicia, y para ello debían cumplir con tener una pensión mayor o igual a la mínima. Por este motivo, este monto fue el que mejor replicó los expuestos del Sistema DL 3.500.

Por otro lado, los datos del IPS no tienen importancia en los primeros años, debido a que se trata de un universo de mayor edad que la del Sistema. Los datos del IPS recién a partir de la edad 80 comienzan a influir para la construcción de las edades extremas.

En el gráfico podemos ver estas comparaciones:



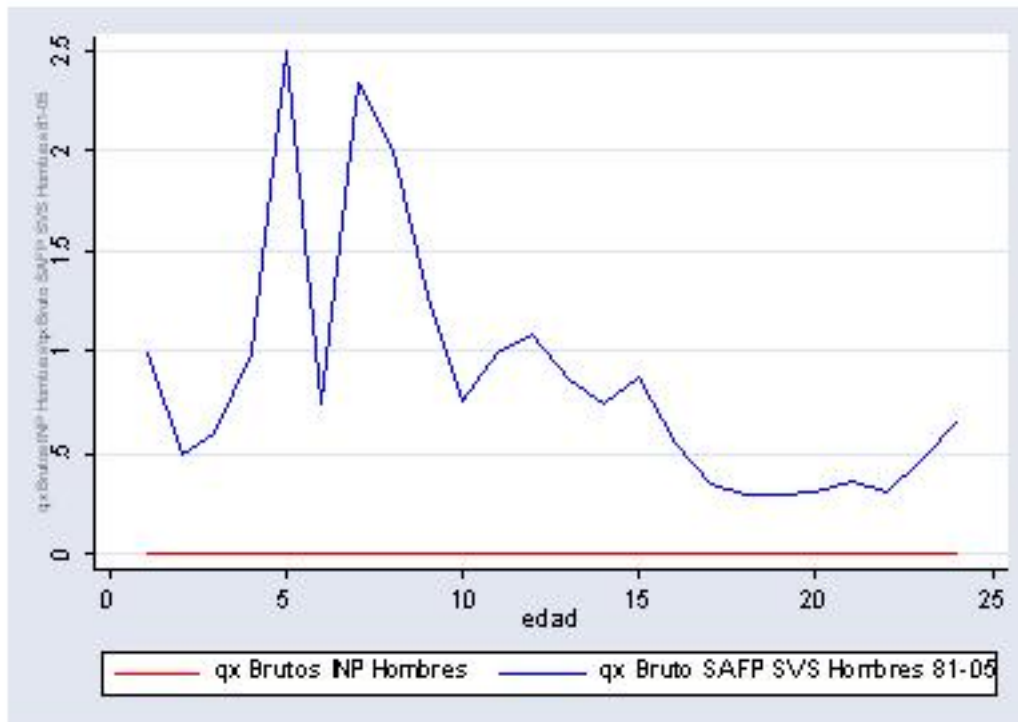
Como podemos ver si hubiéramos tomado la tendencia de la curva verde, esto es utilizando exclusivamente los datos del sistema del D.L. 3.500, tendríamos una caída significativa en la mortalidad.

Si por el contrario, incorporamos el total de los datos del IPS, teniendo en cuenta que a partir de la edad 80 años estos son representativos mientras que los datos del Sistema D.L. N° 3.500 son escasos tendríamos dos efectos, por un lado aumenta la exposición del IPS y por el otro aumenta la mortalidad al incorporar datos de personas de menores ingresos, por lo que se desplazaría hacia arriba la mortalidad del sistema DL 3.500 y cambiaría la tendencia de ésta, la cual es la base de la tabla de mortalidad.

En el peor escenario si sacamos la información del IPS, necesariamente la mortalidad disminuye y se continúa con la tendencia de los datos del sistema, reflejada en la línea verde.

5. Cálculo de expuestos de beneficiarios

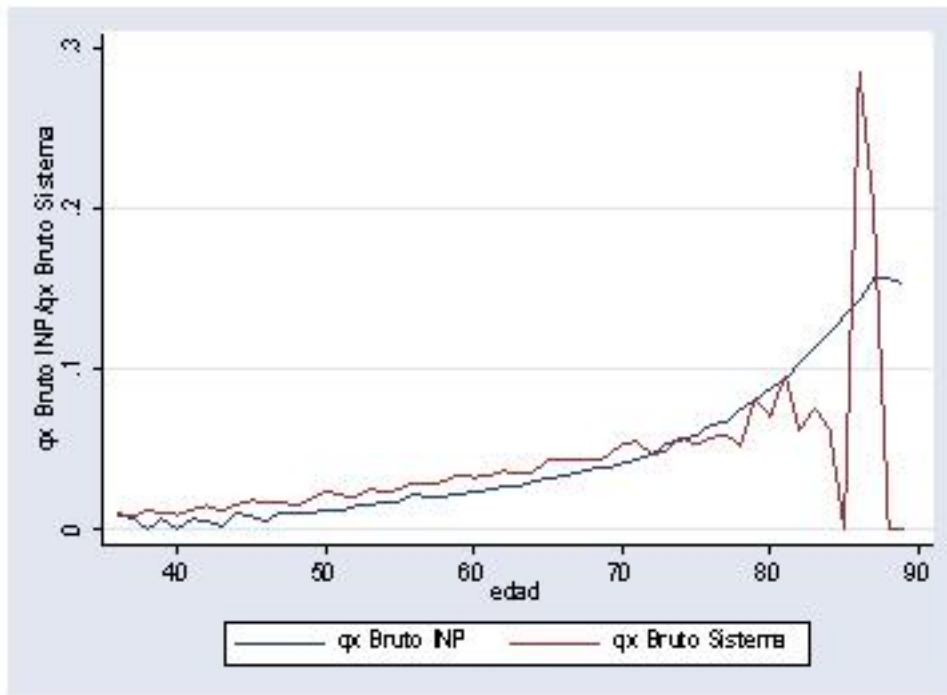
Para el caso de los hombres se analizaron diversos períodos debido a que la información es extremadamente escasa. Incluso con datos desde el año 1995 hasta el 2005, lo que implica utilizar 11 años de observación, no es posible obtener tasas de mortalidad que sean estadísticamente representativas. Podemos ver en el gráfico la gran volatilidad de los datos.



Vemos que en el IPS la información es inexistente, mientras que en los datos del sistema solo se aprecia una tendencia.

6. Cálculo de expuestos de inválidos

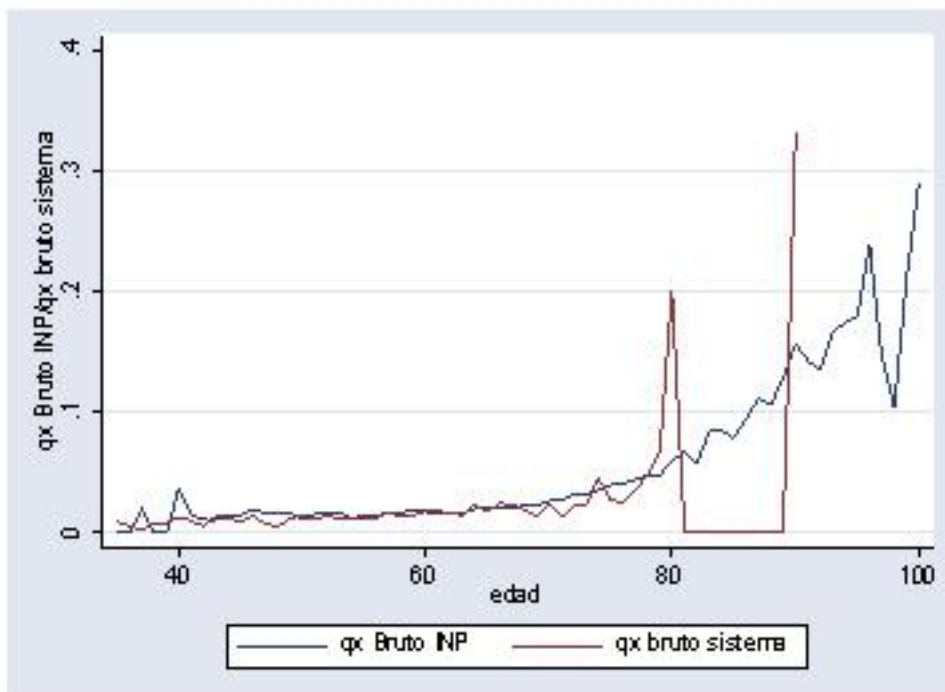
Para el caso de los inválidos se tomaron seis años de exposición desde el 1999 al 2004 inclusive. La información del sistema es escasa por lo que necesariamente se debe complementar con los datos del IPS.



Debido a que existe una diferencia en la mortalidad entre ambos sistemas, que lamentablemente por la escasez de datos no es posible dimensionar, se optó por utilizar los expuestos al riesgo del Sistema desde la edad 30 hasta la 70, los cuales son estadísticamente confiables, y de ahí en adelante los datos del IPS, los cuales solo contribuyen a la construcción de la "cola".

7. Cálculo de expuestos de inválidas

Se consideró el mismo período que para los hombres. En este caso la cantidad de datos es aun menor, pero lo suficientemente confiables como para utilizarlos para la construcción de una tabla.



Se observa que en los primeros años los datos del sistema son representativos y a partir de la edad 65 se utiliza la información del IPS.

IV. Técnicas de ajuste, Wittaker-Henderson Tipo B

Se utilizó, al igual que para la tabla RV-2004, el método Whittaker-Henderson Tipo B para el ajuste. Este método consiste en una combinación de regresión lineal y el método Bayesiano de ajuste.

Podemos definir la fórmula de Whittaker de la siguiente manera:

$$M = F + hS$$

$$F = \sum w_x (q'_{x,t} - q^o_{x,t})^2 \quad S = \sum (\Delta^z q^o_{x,t})^2$$

Donde F (fit) es la medida de ajuste mientras que S (smooth) es una medida de suavidad de la curva. El parámetro h da más o menos intensidad a la suavidad de la curva.

Descripción de F:

Esta es la parte asociada a la minimización de los residuos cuadrados,

$$F = \sum w_x (q'_{x,t} - q^o_{x,t})^2$$

En esta formula el tamaño de la muestra esta ponderando los residuos

$$(q'_{x,t} - q^o_{x,t})$$

Es decir que mientras F tiende a cero el ajuste es mejor.

En los casos en que los residuos tienen una muestra grande (w_x) deben ser más pequeños para mantener F lo más cerca de cero.

El ponderador w_x toma en cuenta la varianza de una distribución normal de la variable aleatoria U_x . Como es sabido U_x es una variable aleatoria binomial pero puede ser aproximada por una variable aleatoria normal siempre que el numero de observaciones n_x se lo suficientemente grande.

$$w_x = 1 / \text{Var}(U_x)$$

Sabemos que la varianza es inversamente proporcional al número de observaciones. A mayores observaciones la varianza disminuye.

$$\text{Var}(U_x) = v_x(1-v_x) / n_x$$

Entonces

$$w_x = n_x / v_x (1-v_x).$$

Aquí se ve claramente que el ponderador da más importancia al q_x bruto cuya varianza es menor (el ponderador va ser más grande).

Descripción de S:

Ahora es necesario definir la medida para definir la suavidad de la curva. Generalmente se utilizan las diferencias finitas de diferente orden sobre los valores de q_x brutos para cuantificar una medida de suavidad.

Esto se puede representar de la siguiente manera:

$$S = \sum (\Delta^z q^o_{x,t})^2$$

"S" se obtiene de la suma cuadrada de las diferencias finitas. Si por ejemplo $z=4$ estamos considerando que la secuencia q_x se asemeja a un polinomio de grado 3. Debemos recordar que las diferencias finitas se asemejan a una derivada y por lo tanto el orden de diferencia condiciona el grado del polinomio.

Para encontrar los valores de q_x ajustados es necesario minimizar M, que es una función con n incógnitas de los valores q_x . Entonces los q_x que minimizan M corresponden a la solución para las n ecuaciones como resultado de la derivada parcial de M con respecto a q_x .

$$\partial M / \partial q_x = 0 \text{ para } x = 1, 2, 3, \dots, N$$

Es posible hallar este resultado representando el problema en forma de matriz.

Coeficientes utilizados para la construcción de las tablas:

Coeficiente	Ajuste Beneficiarias	Ajuste Inválidos	Ajuste Inválidas
H	1.00E+09	1.00E+09	1.00E+10
Diferencia Finita Z	Grado 3	Grado 3	Grado 3

V. Test estadísticos

El principal problema con el ajuste de las tasas brutas de mortalidad es ver si estos realmente representan la población. Para verificar lo anterior se utilizan diferentes test estadísticos que indican que tan fidedigno es el ajuste en comparación con los datos observados.

Los test más comúnmente utilizados son los siguientes:

- **Test de Chi Cuadrado:** Es un test complejo que solo se utiliza como referencia para ver si las tasas brutas de mortalidad representan a dicha población. Mide las desviaciones estándar de la estimación respecto de las tasas brutas de mortalidad. Una de las grandes limitaciones de este test es que para ciertas edades puede haber grandes desviaciones y para otras edades muy pequeñas obteniéndose un resultado final aceptable según el test. Deseablemente estas desviaciones deben ser lo mas constante para obtener un buen resultado. Existen una serie de test adicionales que verifican esta y otras relaciones las cuales son mucho más importantes.

- **Test de Desviaciones Estandarizadas:** Este test estandariza las desviaciones estándar para que sean comparables entre si y fija un nivel de confianza para que estas desviaciones no sean mayor a un cierto numero. Es decir si las desviaciones no son homogéneas y con un grado menor al deseado es muy probablemente que no pase el test.

- **Test de Desviaciones Absolutas:** Este test refleja que las desviaciones absolutas no sean mayores a un cierto número (generalmente se utiliza que éstas no sean mayores a una unidad de la variable normal, o sea de alrededor de 2/3.)

- **Desviaciones Acumuladas:** Los fallecimientos en las distintas edades deben ser independientes, y se deben representar en una variable aleatoria normal. Por esta razón las desviaciones Standard deben ser relativamente homogéneas durante todos los tramos. La hipótesis nula a testear es que las desviaciones acumuladas no deben ser mayores al doble de la raíz cuadrada de la varianza de la distribución.

- **Test de Signos:** Si tomamos en consideración que la hipótesis nula es que las desviaciones observadas de las muertes respecto de las muertes esperadas son una variable normal independiente, lo mismo debe ocurrir con los signos encontrándose un número similar de éstos, tanto positivos como negativos.

- **Test Stevens:** Es similar al test de signos. Estos signos pueden ser todos positivos al principio y luego todos negativos al final del ajuste, por lo que el test de Stevens observa subgrupos de signos a través de la tabla y computa el signo de cada subgrupo. Luego en estos subgrupos deben estar distribuidos de manera similar tanto los signos positivos como negativos.

- **Test Cambio de Signo:** La probabilidad de cada signo es independiente y está representada en una variable normal, por lo que se puede aplicar una variable binomial donde el signo positivo o negativo tiene la misma probabilidad, $\frac{1}{2}$. Con esto se quiere ver que los cambios de signo sean homogéneos y no presenten anomalías.

Resultados obtenidos:

TEST ESTADÍSTICOS

TEST	TABLA		
	BENEFICIARIAS	INVALIDAS	INVALIDOS
Chi cuadrado	☐	☐	☐
Desviaciones Estandarizadas	☐	☐	☐
Desviaciones Absolutas	☐	X	☐
Desviaciones Acumuladas	☐	☐	☐
Test Signo	☐	☐	☐
Test de Stevens	☐	☐	☐
Test Cambio de Signo	☐	☐	☐

VI. Construcción de tablas

1. Construcción de colas

Los datos solo son representativos para un grupo determinado de edades, pero en las edades extremas, mayores a 90 años, es necesario generarlos a partir de algún modelo. Existen diferentes técnicas siendo la más aconsejable utilizar modelos de mortalidad paramétricos para luego extrapolar los resultados.

- Gompertz: $q_x = 1 - e^{-e^{(a+bx)}}$

- Cuadrático: $q_x = 1 - e^{-e^{(a+bx+c)}}$

- Heligman y Pollard: $q_x = ae^{(bx)} / 1 + ae^{(bx)}$

- Kannisto: $q_x = 1 - e^{-(ae^{(bx)} / (1 + ae^{(bx)})) + c}$

a) Caso beneficiarias mujeres:

En este caso el mejor modelo fue el Cuadrático donde se obtuvo el menor error Standard. El problema de este modelo, es que tiende a tener un pendiente más suave, por lo que su forma no representa el caso de las mujeres, donde en esas edades la pendiente debe ser más pronunciada. Es por esta razón que finalmente el modelo de Heligman y Pollard es el más adecuado. Esta pendiente pronunciada se observó también en el caso de las mujeres causantes del sistema previsional chileno.

b) Caso de los beneficiarios hombres:

Debido a que los datos son escasos y no es posible estimar tasas brutas simplemente se optó por utilizar la tabla de causantes hombres RV 2004 y multiplicarla por un factor 1,05. Esta diferencia se observó como promedio entre la tabla de beneficiarias B 06 y las mujeres causantes de la RV 2004.

c) Caso mujeres inválidas:

En este caso se utilizó el modelo de Gompertz, con el que se obtiene una pendiente más pronunciada, y el resultado fue prácticamente el mismo que con los otros modelos, excepto el de Kannisto, cuya pendiente es muy suave.

d) Caso hombres inválidos:

Al igual que en el caso anterior, el modelo de Gompertz fue el que mejor se ajustó, principalmente por su forma. Estadísticamente no había prácticamente diferencia. Por otro lado, nuevamente el modelo de Kannisto fue el más conservador, ya que su pendiente es muy suave.

2. Construcción de edades tempranas

En estos tramos no existen datos para ninguna de las cuatro tablas, por lo que imperiosamente debemos hacer uso de otra fuente.

Para el caso de beneficiarios se optó por la tabla poblacional del INE. El ajuste tendió en forma suave a unirse con la tabla poblacional y se aplicó un factor de conservadurismo del 5%.

Para el caso de los inválidos se observó que el tramo entre 40 y 110 fue prácticamente igual en forma a las tablas anteriores MI 85. Por esta razón se optó por replicar la misma forma de la MI 85 con los factores correspondientes para que sean paralelas.

VII. Factores de mejoramiento y márgenes de seguridad

1. Factores de mejoramiento

Diversos estudios dejan de manifiesto que el mejoramiento de la mortalidad es una realidad sin embargo, no se conoce cual es su límite exactamente. Bajo este escenario se han desarrollado un gran número de modelos que se basan en la experiencia pasada, generalmente más de 25 años de historia. El modelo de capitalización individual es relativamente nuevo, por lo que debemos utilizar la mejora poblacional.

Por otro lado los factores utilizados en el año 2004 corresponden al Censo poblacional del 2002 realizado por el INE (Instituto Nacional de Estadística). Por esta razón ya que no existen nuevos estudios podemos concluir que debemos seguir con los mismos factores de mejoramiento a nivel poblacional utilizados en la tabla de mortalidad RV 2004.

Estas proyecciones fueron implementadas en la tabla de beneficiarios e inválidos de acuerdo a la siguiente fórmula:

$$q_x^{\text{proj}} = q_x^{\text{nuevo}} * (1 - A_x/100)^t$$

donde:

q_x^{proj} : Es el q_x proyectado que refleja el mejoramiento poblacional

q_x^{nuevo} : Son los q_x obtenidos del ajuste de Whittaker Henderson.

A_x : Los factores de mejoramiento

t : Es el número de años desde que las últimas mortalidades fueron proyectadas.

2. Márgenes de seguridad

Se incorporaron márgenes de seguridad para anticipar cualquier cambio brusco en la mortalidad que en el modelo de proyección no haya sido cuantificado.

Estos márgenes de seguridad han sido desarrollados de acuerdo a la siguiente fórmula:

$$\text{Margen} = \frac{\text{Desviación Standard } (U_x)}{q_x^{\circ}} * 100$$

Donde:

q_x° : q_x Bruto

$\text{Desviación Standard } (U_x) = \text{raíz cuadrada } (q_x^{\circ} (1 - q_x^{\circ}) / n_x)$

n_x : Numero de expuestos a edad x

Se utilizó como máximo un margen del 2,5%, tanto para la Tabla MI-2006 y B-2006.