

Informe para Banmédica

**ESTIMACIÓN DE LA VARIACIÓN DE
COSTOS TÉCNICOS DEL GES PARA
EL AJUSTE DE TARIFAS GES
CORRESPONDIENTE AL
PERÍODO 2016-2019**

CLAUDIO SAPELLI



INFORME FINAL

Informe para Banmédica

Estimación de la variación de costos
técnicos del GES para el ajuste de
tarifas GES correspondiente al
período 2016-2019

Claudio Sapelli G.*

**El autor es profesor del Instituto de Economía de la Pontificia Universidad Católica de Chile. Este informe expresa la opinión del autor y sus contenidos no representan la opinión del Instituto de Economía o de la Universidad Católica. El informe se realiza por encargo de Isapre Banmédica. Se agradece la muy eficaz asistencia de Juan Pablo Ossa en la confección del informe.*

Estimación de la variación de costos técnicos del GES para el ajuste de tarifas GES correspondiente al períodos 2016-2019

I.- ANTECEDENTES

Isapre Banmédica S.A. solicitó la colaboración con la División Estudios de la Isapre en el diseño de una metodología que permitiera contribuir a la estimación del índice de reajuste del precio del GES para el periodo 2016-2019. Dicha metodología permitirá la proyección de los costos técnicos que son una parte sustancial del ajuste de precios correspondiente. Este informe entonces, provee la fundamentación para la metodología utilizada. A su vez, sobre la base de los datos públicamente disponibles (complementados por algunos datos entregados por las Isapres) se hicieron en forma independiente las proyecciones. De manera que en este informe también están los resultados que se obtienen con dicha metodología.

II.- METODOLOGÍA DE PROYECCIÓN

II.1.- INTRODUCCIÓN

El objetivo es proyectar las series de tiempo de gastos médicos en las diferentes enfermedades GES y la de beneficiarios, para obtener el cambio en el gasto por beneficiario, que constituirá la base para el ajuste de tarifas que regirá para el periodo 2016-2019.

Se utiliza la metodología de proyección de series de tiempo llamado ARIMA (por sus siglas en inglés, AutoRegressive Integrated Moving-Average). Este método se ha encontrado que es muy bueno para hacer proyecciones, siendo mejor en ello que modelos estructurales de gran complejidad (ver

Referencias). Por lo que a su mejor capacidad predictiva, se suma su sencillez.

Una serie tiempo es una secuencia de observaciones de una variable, ordenados cronológicamente y espaciados entre sí de manera uniforme. Se requiere entonces encontrar la relación que tienen estos datos entre sí, de manera de usar esa información para realizar una proyección.

Esta técnica es abundantemente utilizada en economía y en empresas. Se utiliza por ejemplo para proyectar el empleo y desempleo; para proyectar el precio de un bien en particular; para proyectar ventas; etc. También se utiliza en demografía, para proyectar el número de habitantes, o la tasa de mortalidad. O se puede usar para proyectar la contaminación o la lluvia. En fin, sus usos son muy variados.

El uso del análisis de series temporales para realizar proyecciones se basa en la suposición de que los valores que toma la variable es la consecuencia de tres componentes, cuya actuación conjunta da como resultado los valores medidos.

Estos componentes son:

- 1.- La componente de tendencia.- Este es el cambio a largo plazo del valor de la media de la serie de tiempo considerada. La tendencia se identifica como un movimiento suave de la serie a largo plazo, ya sea creciendo o decreciendo.

- 2.- El componente estacional.- Muchas series temporales presentan cierta periodicidad (aumento de las ventas en torno a fechas determinadas como Navidad, o de determinados productos en

determinadas estaciones del año) o dicho de otro modo, una variación recurrente en cierto periodo (anual, semestral, mensual, etc.). En el caso que nos ocupa como utilizaremos datos anuales este tema no será relevante.

3.- La componente aleatoria.- Esta componente no responde a ningún patrón de comportamiento, sino que es el resultado de factores fortuitos o aleatorios que inciden de forma aislada en una serie de tiempo.

De estos tres componentes los dos primeros son componentes determinísticos, mientras que la última es estocástica.

II.1.1.- Clasificación de las series temporales

Las series temporales se pueden clasificar en:

a.- Estacionarias.- Una serie es estacionaria cuando es estable a lo largo del tiempo, es decir, cuando la media y varianza son constantes en el tiempo. Esto se refleja gráficamente en que los valores de la serie tienden a oscilar alrededor de una media constante y la variabilidad con respecto a esa media también permanece constante en el tiempo.

b.- No estacionarias.- Son series en las cuales la tendencia y/o variabilidad cambian en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

II.1.2.- Autocorrelación

A veces, en una serie de tiempo, los valores que toma una variable en el tiempo no son independientes entre sí, sino que un valor determinado depende de los valores anteriores. Existen dos formas de medir esta dependencia de las variables.

1.- Función de autocorrelación (ACF): La autocorrelación mide la correlación entre dos variables separadas por k periodos.

2.- Función de Autocorrelación Parcial (PACF): La autocorrelación parcial mide la correlación entre dos variables separadas por k periodos cuando no se considera la dependencia creada por los retardos intermedios existentes entre ambas.

II.2.- PROCESOS LINEALES ESTACIONARIOS

II.2.1.- Procesos Autoregresivos

Los modelos autoregresivos se basan en la idea de que el valor actual de la serie, puede explicarse en función de valores pasados. El modelo autoregresivo puede ser de diversos órdenes pero en general son de orden uno (es el valor anterior de la serie que determina el actual), dos (son los dos valores anteriores de la serie los que determinan el actual) o a lo más de orden tres (son los tres valores anteriores de la serie los relevantes). Por ejemplo, en un proceso Autoregresivo de Orden 1 la variable está determinada únicamente por el valor anterior.

Para verificar que el modelo es estacionario para cualquier valor del parámetro, es necesario probar las siguientes condiciones. a) Estacionario en media: Para que el proceso sea estacionario, la media debe ser constante

y finita en el tiempo; y b) Estacionario en covarianza: para que un proceso sea estacionario, la varianza tiene que ser constante y finita en el tiempo.

II.2.2.- Proceso de Medias Móviles

Estos son series de tiempo que muestran un comportamiento “determinado por una fuente externa”. Estos modelos suponen linealidad, el valor actual de la serie está influenciado por los valores de la fuente externa.

Por ejemplo, un proceso de Media Móvil de orden 1: es un modelos de medias móviles en que el valor de la variable se determina en función de la innovación actual y su primer rezago.

A los procesos de medias móviles se les suele denominar procesos de memoria corta, mientras que a los autoregresivos se les denomina procesos de memoria larga.

II.2.3.- Proceso Autoregresivo de Medias Móviles

Es muy probable que una serie de tiempo, tenga características autoregresivos y de medias móviles a la vez. Así, en este proceso habrá términos autoregresivos y términos de media móvil.

Por ejemplo, un Proceso autoregresivo de media móvil de orden (1,1):

Es un proceso en que la variable a pronosticar se determina en función de su pasado hasta el primer rezago, la innovación contemporánea y el pasado de la innovación hasta el rezago 1.

II.3.- PROCESOS LINEALES NO ESTACIONARIOS

Lo anterior se refería a procesos que no tienen tendencia, pero la mayoría de las series que trataremos de pronosticar la tienen, por lo cual lo relevantes es estudiar los procesos para series no estacionarias.

II.3.1.- Proceso Autoregresivo Integrado y de Media Móvil (o ARIMA)

Los modelos de series de tiempo analizados hasta ahora se basan en el supuesto de estacionariedad, esto es, la media y la varianza para una serie de tiempo son constantes en el tiempo y la covarianza es invariante en el tiempo. Pero se sabe que muchas series de tiempo y en especial las series económicas no son estacionarias, porque pueden ir cambiando de nivel en el tiempo o sencillamente la varianza no es constante en el tiempo, a este tipo de proceso se les considera procesos integrados. Por consiguiente, se debe diferenciar una serie de tiempo las veces que sea necesario para hacerla estacionaria y luego aplicarle a esta serie diferenciada el análisis descrito en la sección anterior. En este caso se dice que la serie original es una serie de tiempo autoregresiva integrada de media móvil (o ARIMA, por autorregressive integrated moving average en inglés). Se habla de un ARIMA (p,d,q) en donde p denota el número de términos autoregresivos, d el número de veces que la serie debe ser diferenciada para hacerla estacionaria y q el número de términos de la media móvil.

II.4.- CONSTRUCCIÓN DE LOS MODELOS Y PROYECCIONES

La construcción de los modelos se lleva de manera iterativa mediante un proceso en el que se puede distinguir cuatro etapas:

1. Identificación. Utilizando los datos ordenados cronológicamente se intentará sugerir un modelo que merezca la pena ser investigado. El objetivo es determinar los valores que sean apropiados para reproducir la serie de tiempo. En esta etapa es posible identificar más de un modelo candidato que pueda describir la serie.
2. Estimación. Considerando el modelo apropiado para la serie de tiempo se realiza inferencia sobre los parámetros.
3. Validación. Se realizan contraste de diagnóstico para validar si el modelo seleccionado se ajusta a los datos, si no es así, escoger el próximo modelo candidato y repetir los pasos anteriores.
4. Predicción. Una vez seleccionado el mejor modelo candidato se pueden hacer pronósticos en términos probabilísticos de los valores futuros.

Hoy por hoy los paquetes estadísticos realizan las primeras tres etapas simultáneamente, a través de un comando. En general el modelo identificado es efectivamente el mejor y no puede estimarse otro que se ajuste mejor a los datos o permita hacer predicciones con menor varianza. Solamente en los casos que las series tengan un comportamiento “raro” es que este método fracasa y no identifica ningún proceso. Por ejemplo cuando en los datos hay un outlier, o la serie cambia dramáticamente de comportamiento entre un periodo y otro. En ese caso es necesario eliminar el outlier, o seleccionar el periodo en que uno quiere pronosticar, excluyendo el periodo anterior con comportamiento diferente.

Una de las razones de la popularidad del proceso de construcción de modelos es su éxito en la predicción. Los modelos son buenos para realizar predicciones a corto y mediano plazo.

III.- ANÁLISIS DE LAS PROYECCIONES DEL COSTO TÉCNICO GES PARA BANMÉDICA

Aquí se hace un resumen del procedimiento. Para una descripción paso a paso de lo que se hizo, en que se hace referencia a los archivos de respaldo que acompañan este informe, ver el Anexo 1 de este informe.

Para realizar las proyecciones se hizo un programa para consolidar los gastos en salud según las categorías¹ que se iban a utilizar; se iban a predecir solamente los grupos importantes, y el resto se les iba a proyectar como un todo. Por otro lado al interior de cada grupo se harían proyecciones individuales de los problemas de salud más importantes, tanto en participación como en variación, y nuevamente el resto sería proyectado como un todo (ver cuadro 1 y 2 para un detalle de las series elegidas).

Una vez obtenidos los totales, para el periodo 2010-2015 se armaron los promedios móviles anuales, de tal manera de aumentar la cantidad de datos y la precisión de la identificación del modelo y la minimización del error de la proyección. De esa manera no se tienen datos solo para seis años, sino que para 61 años móviles. La razón de porque se trabaja con años móviles es que, de esta manera, se elimina el problema de la estacionalidad de algunas de las dolencias.

¹ Las categorías a utilizar fueron definidas de acuerdo a los criterios de la SISP, de manera que las agrupaciones utilizadas acá son idénticas a las de la Superintendencia.

Al comparar las bases compiladas por nosotros y la hecha por la Isapre puede observarse que entre ambas no hay diferencias.

Una vez obtenido los datos, se usa el comando “auto.arima” del programa R para que nos ofrezca el modelo que mejor ajusta a los datos. Ese modelo se usa para hacer las proyecciones.

En este caso los modelos seleccionados coinciden en 12 de los 13 casos. Puede observarse en la Tabla 1 que los 13 procesos ARIMA identificados por nosotros para proyectar las series de tiempo de Banmédica son diferentes pero predominan dos tipos de procesos. Por un lado están los caminos aleatorios, con tendencia. Estos son los procesos $(0,1,0)$ con drift; encontramos 6 de ellos. El otro proceso que se encuentra varias veces es aquel con suavizamiento exponencial simple. Estos son los procesos $(0,2,1)$, de los que se encuentran 6.

Es importante aclarar que la razón para proyectar el GES-CAEC como un todo radica en que su comportamiento por patología es muy errático, muy diferente al de la patología que lo activa, y por lo tanto imposible de predecir. Sin embargo el agregado tiene un comportamiento menos errático (como uno esperaría en un seguro).

El proceso en que se encuentra diferencias es en la proyección del GES CAEC de Banmédica (ver anexo 1).

En términos de la proyección de la cartera, nuestra proyección es idéntica a la de la Isapre.

Tabla 1

		Estimaciones ISAPRE	Estimaciones Propias
		ARIMA	ARIMA
Banmedica	B PS18	(2,2,2)	(2,2,2)
	B PS67	(0,2,1)	(0,2,1)
	B RESTO G3	(0,1,0) with drift	(0,1,0) with drift
	B PS4	(0,2,1)	(0,2,1)
	B PS17	(0,2,1)	(0,2,1)
	B RESTO G7	(0,1,0) with drift	(0,1,0) with drift
	B PS15	(0,1,0) with drift	(0,1,0) with drift
	B RESTO G9	(0,2,1)	(0,2,1)
	B PS25	(0,2,1)	(0,2,1)
	B PS37	(0,1,0) with drift	(0,1,0) with drift
	B RESTO G4	(0,1,0) with drift	(0,1,0) with drift
	RESTO G B	(0,1,0) with drift	(0,1,0) with drift
	B GES-CAEC	(2,2,2)	(0,2,1)

Conjugando la proyección de los gastos, sumados, y dividido este gasto total por la proyección de beneficiarios, nos da lo que puede verse en la Tabla 2.

Tabla 2

	Banmédica	
	Estimaciones ISAPRE	Estimaciones Propias
Total por Beneficiario Año 2015	\$ 5.507	\$ 5.506
Total por Beneficiario Año 2018	\$ 8.081	\$ 8.106
Crec. Punto Medio 12-15, Punto Medio 15-18	46,1%	46,3%
Crecimiento 2015-2018	46,8%	47,2%

La proyección del costo técnico para Banmédica nos da un aumento de 47,2% en tres años (si bien la tarifa rige desde el 7/16 al 6/19 por falta de datos acá se hizo la proyección desde fines del 2015 hasta fines del 2018). El ajuste tarifario tiene que estar centrado en la mitad del periodo, por lo que corresponde calcular el gasto en el punto medio y después obtener el crecimiento entre ese gasto y el del punto medio del periodo anterior (2013-2016). Ese es el ajuste del costo técnico que tiene que ser incorporado en el

cambio en la tarifa GES. Este número da 46,3% para Banmédica, un número casi idéntico al estimado por la Isapre.

IV.- CONCLUSIONES

Se realizaron proyecciones de los gastos médicos de los problemas de salud incluidos en el GES para Banmédica. Ellos se realizaron por el método considerado óptimo, por su mezcla de sencillez y eficacia, el método ARIMA. En este informe se describe el método y se realizan las proyecciones en forma independiente, encontrando resultados (46,3%) que son similares a los encontrados por la Isapre (46,1%).

REFERENCIAS

Box, George; Jenkins, Gwilym (1976), "*Time Series Analysis: forecasting and control*", rev. ed., Oakland, California: Holden-Day

Cowpertwait P.S.P., Metcalfe A.V. (2009), "*Introductory Time Series with R*", Springer.

Durbin J., Koopman S.J. (2001), "*Time Series Analysis by State Space Methods*", Oxford University Press.

Gershenfeld, Neil (2000), "*The Nature of Mathematical Modeling*", Cambridge University Press.

Hamilton, James (1994), "*Time Series Analysis*", Princeton University Press.

Shasha, D. (2004), "*High Performance Discovery in Time Series*", Springer.

Shumway R. H., Stoffer (2011), "*Time Series Analysis and its Applications*", Springer.

Woodward, W. A., Gray, H. L. & Elliott, A. C. (2012), "*Applied Time Series Analysis*", CRC Press.

ANEXO 1: Detalle de Procedimientos y Estimaciones

Para revisar las estimaciones realizadas por el equipo de la Isapre Banmédica se realizó lo siguiente:

1.- Primero se recibió la base de datos actualizada con la información necesaria de los procedimientos, montos y problemas de salud asociados a todos los gastos GES y GES-CAEC de ambas Isapres. Esta base es una corrección de la base que entregan a la Superintendencia con la diferencia de que tiene en muchos procedimientos mejor asociado el problema de salud al que corresponde (internamente cambiaron valores “sin problemas de salud” por los que correspondían). Esta base se encuentra contenida en la carpeta “Agrupaciones STATA” y tiene el nombre de “BASE_ISAPRES.dta”.

2.- Se procedió a agrupar (sumar) los gastos en GES de cada Isapre (VALOR_BONIFICADO) por problema de salud y fecha en que se realizó. Luego se realiza el mismo procedimiento para el GES-CAEC con la diferencia de que este se suma sólo por fecha para cada Isapre. El archivo (do file) que realiza este colapso de la base de datos se encuentra en la carpeta “Agrupaciones STATA” y tiene el nombre de “Creación Agrupaciones”. Al correr el archivo en STATA se genera la base (en un archivo Excel) a utilizar que recibe el nombre de “Base_General”.

3.- Posteriormente, se comienza a trabajar en Excel. El archivo a utilizar es “Comparación Estimaciones”. La primera pestaña de este archivo tiene el nombre de “Base Creada en Stata” que es la importación de “Base_General”, luego se realiza una agrupación por años móviles que serán las series a proyectar. La segunda pestaña (“Comparación Bases”) tiene en la parte superior la agrupación que generaron en el grupo de investigación de las Isapres, y en la parte inferior la agrupación que proviene de la “Base Creada en Stata”. Estas agrupaciones se realizan según los grupos de problemas de salud definidos para el GES y se agrupan además individualmente problemas de salud que tienen un gasto importante dentro del grupo, el detalle de los criterios de agrupación de problemas de salud y series a estimar se pueden ver en la figura al final del

documento. En la comparación de la base las series a proyectar no hay diferencia, ninguna difiere en más de un 0,0% con respecto a la estimada por las Isapres.

4.- Luego de generadas las 13 series a proyectar, estas se exportan en un archivo .txt para operarlas en R. El archivo que busca el proceso ARIMA que mejor se ajuste a los datos y luego realiza las proyecciones para cada serie se llama "ARIMA1.R". Desde este archivo se van extrayendo y pasando al Excel "Comparación Estimaciones" en la pestaña "Comparación Proyecciones". En la parte superior de esta pestaña se muestran las proyecciones realizadas por las Isapres y en la parte inferior las que se realizaron en el archivo "ARIMA1.R". Sólo se encuentra una divergencia en las proyecciones:

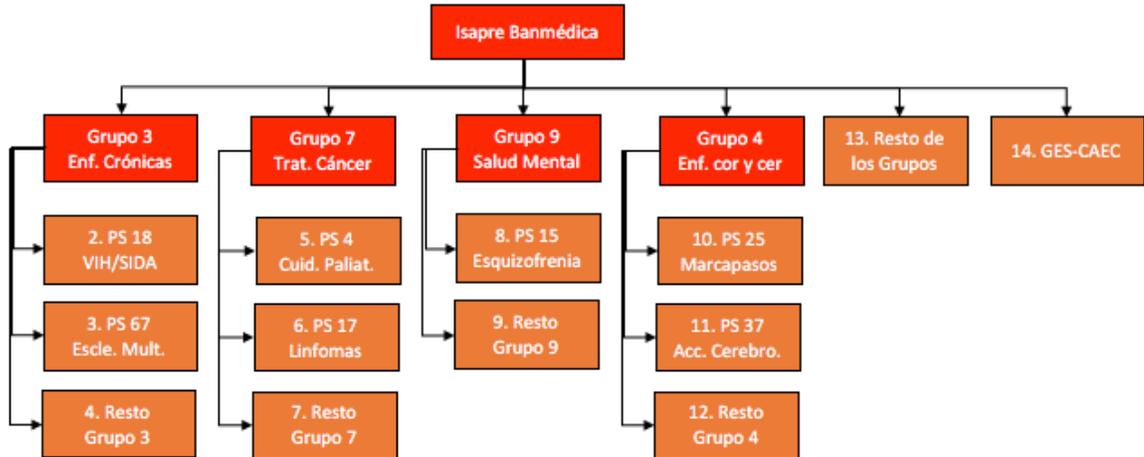
En la serie que agrupa al GES-CAEC encontramos un proceso ARIMA (0,2,1) que difiere del (2,2,2) reportado por la Isapre. En este caso el proceso (0,2,1) se ajusta mejor a los datos y proyecta valores que llegan a ser hasta un 2,9% mayores al que proyecta el ARIMA reportado por las Isapres.

En el extremo derecho de la pestaña "Comparación Proyecciones" se muestran las series sumadas. La divergencia encontrada con lo proyectado originalmente por la Isapre es del orden del 0,1% al 0,3% (mayores nuestras estimaciones a las originales).

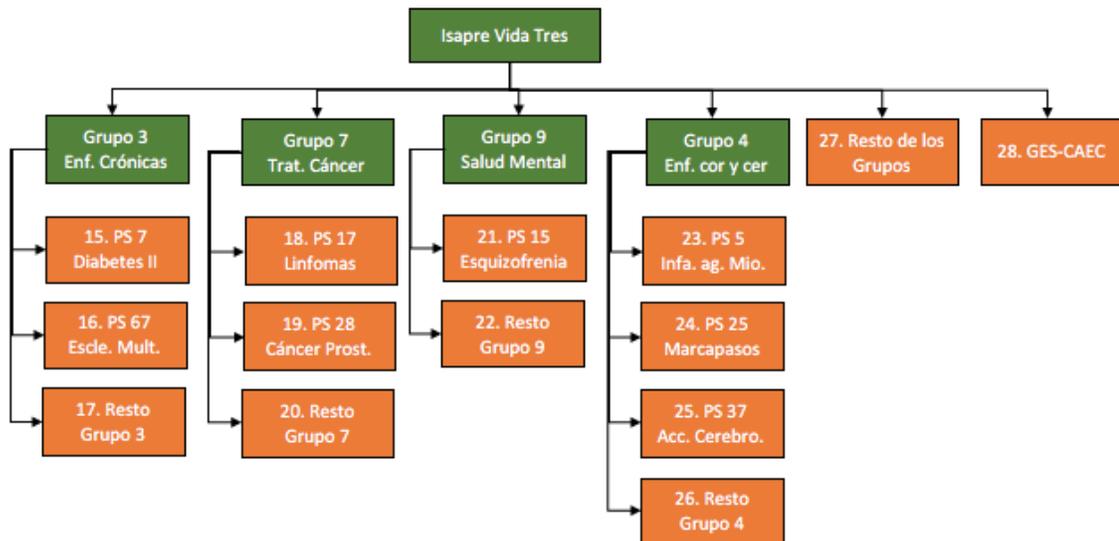
5.- Por último se estima un proceso ARIMA para la cartera, este también se extrae de correr el archivo "ARIMA1.R". Aquí no se encuentran diferencias con lo estimado originalmente por las Isapres. Por lo que el gasto por beneficiario (original versus proyectado en este documento) varía en exactamente en los mismos porcentajes que los gastos agregados por Isapre.

Elección de series a proyectar:

Cuadro 1



Cuadro 2



ANEXO 2: Bases de datos entregadas

Se deja expresa constancia que las bases de datos entregadas por Isapre Banmédica S.A. al Prestador, son fiel reflejo de la normativa de la Superintendencia de Salud en materia de Archivos Maestros de Prestaciones, Archivos Maestros de Beneficiarios, y Archivos de Solicitudes GES, sin perjuicio del resguardo de la información acerca de la identificación de los beneficiarios.

Archivos Maestros	Período
Prestaciones_Bonificadas	201201 al 201512
Beneficiarios	201201 al 201512
Diccionario/Descripción Códigos	
Solicitudes GES	201201 al 201512

Luego, se realizó una segunda entrega de datos. Las Isapres a través de procesos internos pudieron mejorar la información, en particular respecto de la clasificación de los problemas de salud a que estaban asociados procedimientos y medicamentos. Es con esta segunda base corregida que se realizaron las estimaciones.