UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

Programación Lógica Orientada a Riesgos Actuariales

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARÍA

P R E S E N T A:

LAIMA MICHELLE CID LÓPEZ

DIRECTOR DE TESIS:
ACT. HARIM GARCÍA LAMONT
2015

Ciudad Universitaria, D. F.

1. Datos del Alumno

Cid
López
Laima Michelle
55 93 31 50
Universidad Nacional Autónoma de México
Facultad de Ciencias
Actuaría
305348388

2. Datos del Tutor

Act.
Harim
García
Lamont

3. Datos del sinodal 1

Dr.
Yuri
Salazar
Flores

4. Datos del sinodal 2

M. en C.
Fernando Daniel
Pérez
Arriaga

5. Datos del sinodal 3

M. en C.
Óscar
Ortega
Ibañez

6. Datos del sinodal 4

Act.
Alberto
López
Enríquez

7. Datos de la tesis

Programación Lógica Orientada a Riesgos Actuariales
47 p.
2015

Dedicatoria y Agradecimientos

A mi familia
Le dedico esta tesis ya que han estado a mi lado apoyándome y guiándome durante todo este tiempo del proceso de realización del trabajo dándome ánimo, acompañándome en los momentos de crisis y en los momentos de felicidad. A mis padres por su comprensión y ánimos para salir adelante, para poder por fin lograr la culminación de este proyecto, que además es la mejor herencia que me pueden dar.

Agradezco tanto a mi familia como amigos, que estuvieron ahí para apoyarme, su paciencia y tiempo que me han brindado, ya que gracias a ustedes, su apoyo y consejos he podido llegar a donde estoy, cerrando una etapa dentro de mi vida personal y profesional.

A mi director de tesis Act. Harim García Lamont
Le agradezco por su tiempo, paciencia y enseñanzas indispensables para la elaboración y culminación de este proyecto de tesis

A la Universidad Nacional Autónoma de México
Por haberme abierto sus puertas para poder llevar a cabo mis estudios profesionales.

Índice

I.- Introducción

En la actualidad el proceso de suscripción de una solicitud de seguros de vida varía según la empresa, pero existen criterios comunes utilizados por las compañías de seguro. La suscripción le permite a las empresas clasificar riesgos y cotizar adecuadamente el costo del seguro para cada solicitante. Determina la cantidad del seguro que se extiende a las personas, de tal manera que la compañía de seguros pueda todavía obtener un beneficio. El proceso de suscripción pronostica si eres un cliente de alto o bajo riesgo[12].

En general, no existe un proceso uniforme para la suscripción en la industria de los seguros. Como resultado, cada compañía desarrolla sus propias políticas para tomar las decisiones, siendo objetivas en sus análisis de riesgo, y para mejorar aún más dicha objetividad se ayudan de la tecnología, es decir de los programas que eficientan dicho proceso.

Dando pie al desarrollo del presente trabajo, comenzando con una reseña de la cultura en la administración de riesgos, el proceso para la administración de riesgos, y el contexto de los riesgos para una institución aseguradora al generar las reservas por medio del pago de las primas e inversión en instrumento financieros.

En el capítulo 2 se establece la perspectiva de riesgo para una institución aseguradora en la generación de reservas contemplando únicamente el modelo actuarial, por lo que se establece el riesgo asociado a una póliza de seguro (Siniestralidad); el cálculo de la reserva mediante el cálculo actuarial. Posteriormente, mediante la inversión de la reserva en un instrumento financiero libre de riesgo, se intenta incrementar el valor de la misma considerando el rendimiento asociado, por lo que se establece el modelo financiero y el riesgo correspondiente.

En el capítulo 3 se establece un esquema cognitivo de los riesgos ya mencionados y considerando factores de riesgo. Para esto se establece una cadena de Markov oculta la cual ayuda a modelar la dinámica entre los riesgos y sus factores, y una herramienta para establecer tendencias en los riesgos, sus probabilidades y la toma de decisiones.

En el capitulo 4 se hace la estimación tanto para el modelo actuarial como para el financiero y la corrida del modelo Markoviano con respecto a la información de los datos. Haciendo uso de: Análisis de Supervivencia y programación dinámica.

En el capítulo 5 se da a conocer la síntesis y las conclusiones del trabajo.

II.- Antecedentes

En el capítulo se da a conocer a grandes rasgos el proceso de una compañía de seguros para poder plantear el concepto de Riesgo y así llegar a la definición que se tiene en la perspectiva actuarial y en la financiera, además de dar a conocer el proceso de la Administración de Riesgos.

Cultura para la Administración de Riesgos

La administración de riesgos es una de las disciplinas que actualmente se está desarrollando, y que depende de la perspectiva donde se aplica, ya sea comercial, industrial o tecnológico. Como consecuencia se han desarrollado diversas metodologías y estándares, por lo que cada perspectiva puede diferir en ciertos conceptos y caracterizaciones del riesgo. Por ejemplo en la industria aseguradora la administración de riesgos va vinculada en la generación de planes de contingencia y compensaciones económicas, mientras que en ingeniería la administración de riesgos podría ser aplicada para la seguridad en procesos de manufactura.

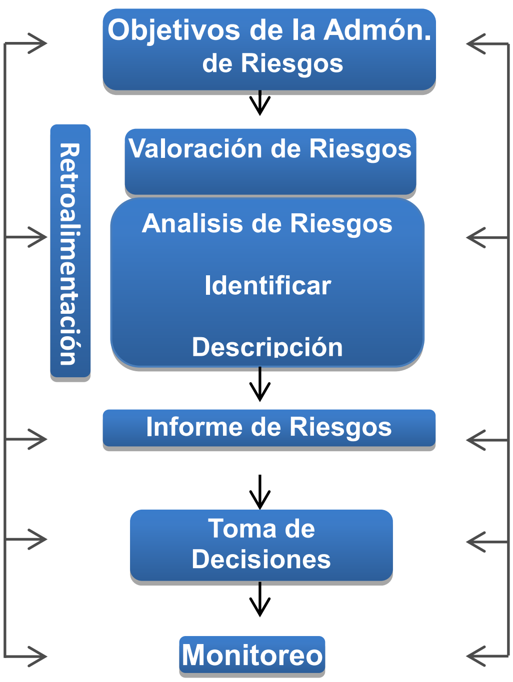
Así entonces para poder entender mejor lo que es el riesgo daremos a conocer algunas de las definiciones que existen:

- Por riesgo se entiende la probabilidad de que se desencadene un determinado fenómeno o suceso que, como consecuencia de su propia naturaleza o intensidad y la vulnerabilidad de los elementos expuestos puede producir efectos perjudiciales en las personas o pérdidas de bienes [13].
- Es la probabilidad de que suceda un evento, impacto o consecuencia adversos. Se entiende también como la medida de la posibilidad y magnitud de los impactos adversos, siendo la consecuencia del peligro, y está en relación con la frecuencia con que se presente el evento [3].
- Es una medida de potencial de pérdida económica o lesión en términos de la probabilidad de ocurrencia de un evento no deseado junto con la magnitud de las consecuencias [17].
- Riesgo es la vulnerabilidad ante un posible potencial de perjuicio o daño para las unidades o personas, organizaciones o entidades. Cuanto mayor es la vulnerabilidad mayor es el riesgo, pero cuanto más factible es el perjuicio o daño, mayor es el peligro. Por tanto, el riesgo se refiere sólo a la teórica "posibilidad de daño" bajo determinadas circunstancias, mientras que el peligro se refiere sólo a la teórica "probabilidad de daño" bajo esas circunstancias [14].

- Es la posibilidad de pérdida o daño. El hombre desde que nace vive con la constante amenaza de enfermedad, accidente, muerte... De la misma forma sus propiedades pueden sufrir incendios, robos, etc. [6]
- Es tener el monto necesario en reserva para enfrentar las posibles reclamaciones que se puedan presentar en el futuro en caso de ocurrir el siniestro [16].

Las clasificaciones más representativas del riesgo son: *Puros vs Especulativos* y *Objetivos vs Subjetivos*. Los *Riesgos Puros* son aquellos en los que únicamente existe un impacto adverso (pérdida) y los *Riesgos Especulativos* son aquellos en los que se contemplan ambos *impactos*; *adversos y favorables (ganancias)*. Además de las perspectivas existen otros conceptos relevantes como son: las *Fuentes y Factores de Riesgo*. Las *Fuentes de Riesgo* son las causales de un impacto adverso o favorable en un agente, y los *Factores de Riesgo* son las causales de una o varias Fuentes de Riesgo.

Dicho lo anterior la Administración de Riesgos ha sido definida por muchas organizaciones basándose en lo que es un riesgo para ellos y las reglas o criterios que se deben seguir para su administración. Como resultado del conocimiento técnico se tiene “El proceso para la Administración de Riesgos”. El siguiente esquema representa los pasos dentro del proceso para la administración de riesgos.



Proceso para la administración de Riesgos

Como se menciona en el ©AIRMIC, ALARM, IRM: 2002, translation copyright FERMA: 2003[3]

Proceso para la administración de Riesgos [3]

1) **IDENTIFICACIÓN DEL RIESGO:** Determinar cuáles son las exposiciones más importantes al riesgo en la unidad de análisis (Organización).
2) **VALORACIÓN DEL RIESGO:** Es la cuantificación de los costos asociados a riesgos que ya han sido identificados.
3) **Análisis de riesgos**: Comprende la identificación, descripción y estimación de riesgos.

    3.1) Identificación de riesgos, propone identificar la exposición de una empresa a la incertidumbre. Ello requiere un conocimiento detallado de dicha empresa, así como el desarrollo de una visión común coherente de su estrategia y de sus objetivos operacionales, incluyendo los factores críticos para su éxito y las amenazas y oportunidades relacionadas con la consecución de estos objetivos.

    3.2) Descripción de riesgos, el objetivo de la descripción de riesgos es mostrar los riesgos identificados de una forma estructurada, por ejemplo, utilizando una tabla como se muestra a continuación.

Nombre del riesgo	*Ej. Riesgo Financiero.*
Alcance del riesgo	*Descripción cualitativa de los sucesos, su tamaño, tipo, número y dependencias.*
Naturaleza del riesgo	*Ej. Estratégicos, operacionales, financieros, de gestión del conocimiento y de conformidad.*
Interesados	*Interesados y sus expectativas*
Cuantificación del riesgo	*Importancia y probabilidad*
Tolerancia del riesgo / Apetito	*Potencial de pérdida e impacto financiero del riesgo* *Valor en riesgo* *Probabilidad y tamaño de las pérdidas/ganancias potenciales* *Objetivo(s) del control de riesgo y nivel deseado de rendimiento*
Tratamiento del riesgo y mecanismos de control	*Medios primarios por los que se gestiona el riesgo actualmente* *Niveles de confianza en el control existente* *Identificación de protocolos de supervisión y revisión*
Acción potencial de mejora	*Recomendaciones para reducir riesgos*
Política y estrategia a desarrollar	*Identificación del responsable de la función de desarrollo de la política y la estrategia.*

Tabla - Descripción de riesgos[3]

El uso de una estructura bien diseñada nos ayuda asegurar un proceso exhaustivo de identificación, descripción y valoración de riesgos. Dándonos la consecuencia y probabilidad de cada uno de los riesgos, con lo que será posible dar prioridad a los riesgos clave para analizarlos con más detalle.

3.3) La estimación de riesgos puede ser cuantitativa, semi-cuantitativa o cualitativa en términos de probabilidad de ocurrencia y de sus posibles consecuencias.

4) **Integración de riesgos,** es para poder obtener un ordenamiento de los riesgos involucrados en la organización, considerando los factores para la constitución de la reserva.
5) **Informe de Riesgos**, con este se obtienen las consecuencias que se pueden ver en términos de amenazas (riesgos puros) y oportunidades (riesgos especulativos) las cuales a su vez pueden dividirse en altas, medias o bajas.
6) **Toma de Decisiones,** depende de la postura que se quiera tomar;
   - *Evitar el riesgo* (no exponerse a un riesgo determinado);
   - *Prevención y control de pérdidas* (medidas tendientes a disminuir la probabilidad o gravedad de pérdida);
   - *Retención del riesgo* (absorber el riesgo y cubrir las pérdidas con los propios recursos) y finalmente,
   - *La transferencia del riesgo* (que consiste en trasladar el riesgo a otros, ya sea vendiendo el activo riesgoso o comprando una póliza de seguros).
7) **Monitoreo:** Las decisiones se deben de evaluar y revisar periódicamente.

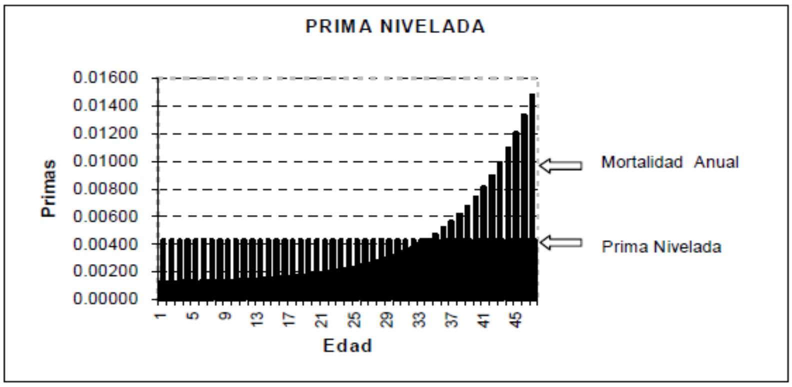
Por otro lado, las organizaciones buscan automatizar el proceso para la administración de riesgos. En la actualidad el concepto BigData establece infraestructuras para el proceso de almacenamiento, procesamiento y análisis de datos.

Dinámica de la compañía

Para una institución de seguros uno de los aspectos básicos de la regulación y supervisión de las operaciones de seguros se basa en lograr que las instituciones cumplan con las obligaciones que han contraído con los asegurados. El cumplimiento de tales obligaciones consiste fundamentalmente en hacer frente a las reclamaciones futuras que se presenten, para lo cual las instituciones deben contar con los recursos financieros suficientes. El principal recurso con que cuenta una aseguradora para tales efectos son las reservas, por lo que es fundamental establecer criterios generales para su constitución en las instituciones de seguros.

La Reserva Matemática: Se refiere al monto que se tiene correspondiente a los seguros de vida, se generan por medio de la Prima de Riesgo que corresponde al costo esperado de la siniestralidad, es decir es la porción de la prima que debe destinarse al pago de las reclamaciones por concepto de siniestros[16].

Particularmente en los seguros de vida, la constitución de la reserva matemática se realiza dependiendo de la temporalidad del seguro, es decir, el tiempo que dura la cobertura del seguro, y la forma del pago de la prima. Para los seguros de vida con temporalidad mayor a un año, es frecuente que el pago de las primas se haga en forma nivelada y anual. La forma de operación de estos seguros origina la necesidad de constituir una reserva, ya que la prima nivelada anual al principio del tiempo es superior a la mortalidad esperada, y a partir de cierto número de años transcurridos, esta prima es inferior a la mortalidad esperada anual. Lo anterior se debe a que el riesgo de muerte es creciente con la edad de los asegurados mientras que la prima nivelada, al ser un valor promedio, no corresponde al valor esperado de la mortalidad anual.



Como se aprecia en el gráfico, al principio del tiempo, la prima nivelada que paga el asegurado es superior a la cantidad que debería pagar, por lo que existe un exceso llamado prima de ahorro. Esta prima de ahorro debe ser reservada para años futuros cuando esta situación se invierta y la prima nivelada resulte insuficiente para el pago de la siniestralidad esperada anual.

Otro aspecto importante de las reservas de este tipo de planes, es que en muchas ocasiones se efectúan altos gastos de adquisición en los primeros años. Estos principalmente se derivan de la contratación del producto, es decir, a la publicidad y comisiones pagadas a los agentes, por lo que se presentan eventuales pérdidas técnicas.

En estos casos, existen procedimientos que mediante la modificación del método tradicional de constitución de la reserva, permiten a la institución constituir una reserva inferior y disponer, en los primeros años, de mayores recursos que permitan compensar las pérdidas. A los citados métodos se les conoce con el nombre de Sistemas Modificados de Reserva.

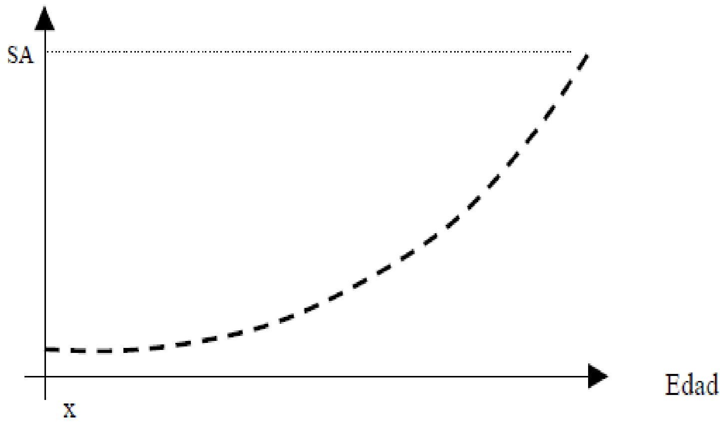
Si se trata de un seguro de vida entera para una persona de edad x, con pago de prima nivelado, la reserva matemática en el año t, se simboliza en la notación actuarial como ${}_tV_x$ y debe calcularse como:

$${}_tV_x = A_{x+t} - PN_x * \ddot{a}_{x+t}$$

dónde:

A_{x+t} representa el valor presente actuarial de las obligaciones futuras de la compañía, por concepto de siniestros futuros[16]

$PN_x * \ddot{a}_{x+t}$ representa el valor presente actuarial de las obligaciones futuras del asegurado, por concepto de pago de primas.



En el gráfico se puede apreciar que la reserva matemática debe aproximarse al valor de la suma asegurada (SA) contratada, a medida que la persona tiene mayor edad, lo que es consecuencia del aumento del riesgo de muerte por envejecimiento.

Aunque las reservas matemáticas son constituidas para enfrentar las obligaciones futuras por concepto de siniestros, su utilización no se realiza disponiendo de la reserva en forma directa. La reserva se va ajustando gradualmente con el tiempo mediante un cálculo periódico (anual) y su saldo puede resultar en un incremento o decremento dependiendo del momento en que se encuentre la vigencia del plan. Por lo que las reclamaciones por concepto de siniestros del año deben pagarse con los diferenciales que resulten entre las primas cobradas y el incremento o decremento de la reserva matemática.

En caso de que la parte devengada de las primas no resulte suficiente para el pago de siniestros, la institución deberá responder con sus propios recursos de capital, lo cual se identifica como una situación de pérdida.

Por otra parte , la *Reserva para Fluctuación de Inversiones* tiene como objeto, enfrentar la posible pérdida que se produce cuando, por efecto de fluctuaciones y situaciones imprevistas en los mercados financieros, no se logran obtener los rendimientos mínimos necesarios para obtener ganancias en la empresa, o se producen pérdidas por la inversión en ciertos instrumentos financieros.

La constitución de esta reserva se realiza con aportaciones periódicas que se derivan de la utilidad en los productos financieros de la inversión de las reservas matemáticas. Su límite quedará definido por una cantidad que corresponda a la pérdida esperada anual por una fluctuación imprevista en los instrumentos financieros del mercado de inversión.

Es necesario señalar que toda reserva debe tener como objeto compensar en el tiempo, los quebrantos financieros que pueda tener una compañía de seguros por circunstancias o eventos imprevistos, o para el cumplimiento de obligaciones futuras que se derivarán de la operación directa del seguro. Por lo que se quiere tener un mejor rendimiento en las inversiones de las reservas (riesgo de mercado).

De lo anterior se establecen los riesgos asociados en la institución de seguros: Riesgo Actuarial y Riesgo Financiero.

Perspectiva de Riesgo Actuarial

El riesgo actuarial está determinado por el fallecimiento o la supervivencia del asegurado, que tiene asociado diversos factores de riesgo (edad, sexo), que se miden utilizando las tablas de supervivencia o de mortalidad. La manera en que se cubre este riesgo es por medio de un contrato de seguros. En este la compañía o entidad aseguradora se compromete a cubrir económicamente la pérdida o daño que el asegurado pueda sufrir durante la vigencia de dicho contrato a cambio de que el asegurado cumpla con la obligación de pagar una prima. Esta prima está relacionada con la esperanza de vida y los gastos de adquisición y administración del seguro.

El siguiente diagrama sintetiza la relación entre fuentes y factores de riesgo sin considerar el tiempo.

Esquema de Descripción del Riesgo

Perspectiva de Riesgo Financiero

Cuando un seguro es a largo plazo, habitualmente existe otro riesgo que asume la aseguradora, el riesgo financiero. De esta manera la compañía debe pagar una determinada suma asegurada dado el riesgo actuarial, es decir, las reclamaciones futuras de los asegurados. Por lo tanto la aseguradora debe invertir, de tal forma que el rendimiento generado por el instrumento cumpla con el monto, que tenga liquidez y un plazo adaptado a los tiempos esperados de supervivencia.

Dado lo anterior se tiene como objetivo la construcción de un sistema para la administración de riesgos. A continuación se enlistan los pasos para dicho sistema:

1.- Participación del Cliente en el proceso como agente; proporcionando sus datos a la compañía aseguradora (edad y tipo de habito si fuma o no).

2.- Envío de la información al grupo de análisis técnico para detectar los riesgos, usando aplicaciones de la aseguradora (sistema de administración de riesgos).

3.- Verificación del riesgo con la red bayesiana (cadenas de Markov ocultas), para clasificar al cliente.

4.- Obtención del modelo de riesgos final que contempla la prima y la reserva para las contingencias. Calculo de Primas y Reserva de la Nota Técnica del producto de la institución.

$$PA = \frac{PU}{\ddot{a}_{x:w-x}} \text{ y } {}_tV_x^{min} = {}_tV_x - AM_t$$

5.1.- El cliente recibe la respuesta del seguro, se acepta o se rechaza el cliente.

5.2.- Si es aceptado, el cliente empieza el pago de primas y la generación de la reserva. Como se mencionó en los sistemas modificados de reserva se utilizara: la perdida amortizable para el primer año, que será el mínimo entre la Perdida Esperada del primer año y la Prima de Ahorro (diferencia entre la prima neta nivelada y el costo de siniestralidad esperado del primer año) y la reserva para los siguientes años.

$$PA_1 = mín(PE_1, PAh_1) \text{ y } {}_tV_x$$

6.- Se envían las características técnicas del modelo al Dpto. de sistemas informáticos para el control y monitoreo del cliente.

7.- Se verifica al asegurado y se invierte de manera adecuada reserva, para cumplir la directiva del proceso propuesto. Se genera la reserva y el rendimiento.

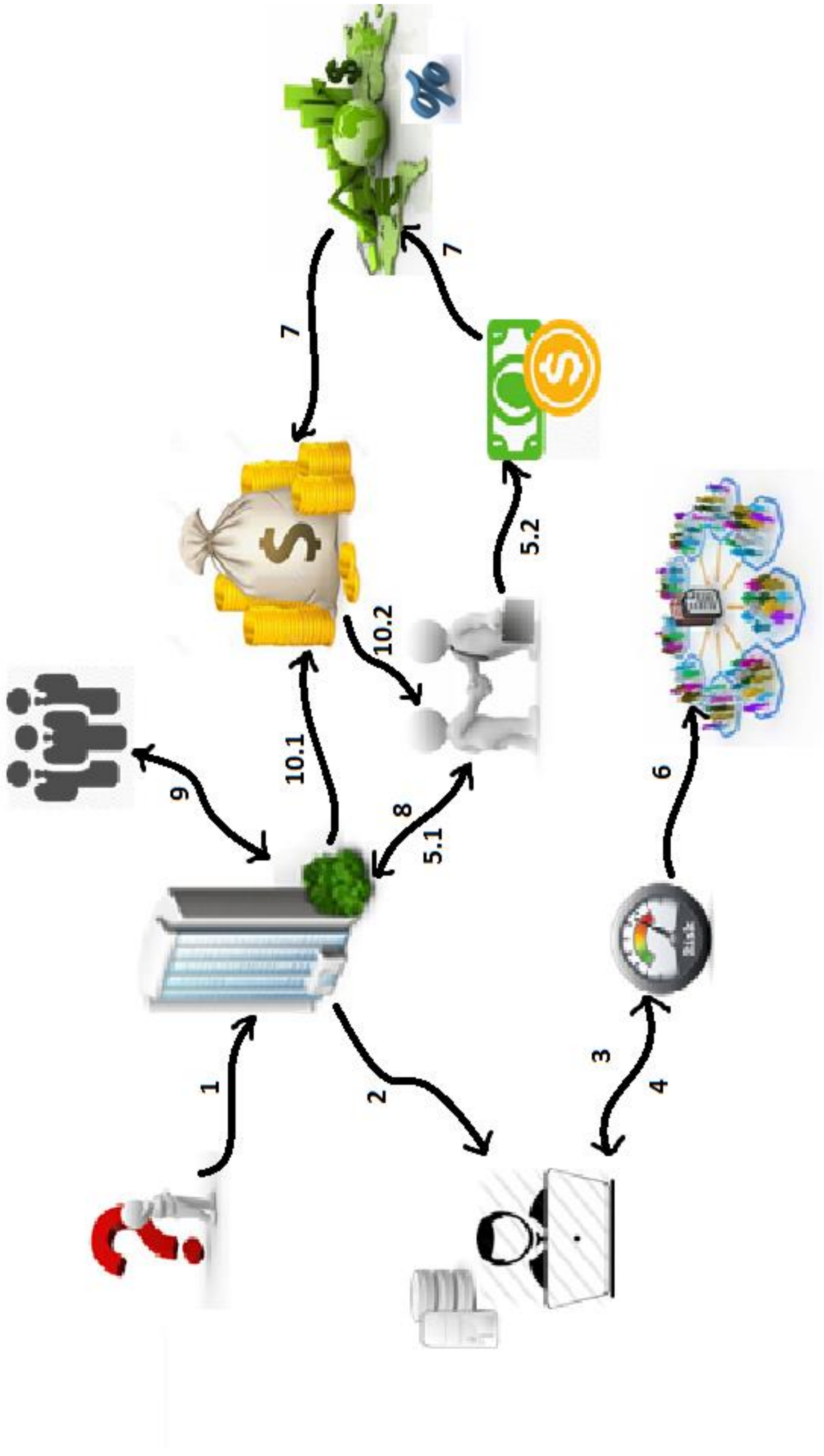
8.- En caso de siniestro, se envía la información correspondiente a la institución para su procesamiento.

9.- La información es procesada para su peritaje con un grupo de abogados y contadores. Sin evidencia de fraude, se acepta el pago de la contingencia.

10.1.- La aseguradora accede a caja.

10.2.- Se hace el pago del siniestro.

Diagrama de Flujo



III.- Modelo Asociado

*E*ste capítulo se constituye por la descripción del modelo con relación a las fuentes y factores de riesgo asociados, además del tiempo.

Información asociada

El análisis que presento a continuación lo realice con la información de pólizas que participaron en un periodo de 5 años de un producto que más adelante se describe. Para dicho análisis observe el comportamiento de las reclamaciones para obtener la función de riesgo y la función de supervivencia de la cartera. Esto con la finalidad de conocer cuál es el tiempo optimo en el que se puede invertir de forma segura la reserva y generar una mayor utilidad a la compañía. La base de la institución contiene Suma Asegurada, Prima Emitida, Reserva, Edad Real, Sexo, Tipo de hábito, duración. La información se utilizó en partes, pues para el análisis descriptico se usó la edad y tipo de habito, ya que el sexo no estaba completo, y tanto la Suma asegura como la prima emitida y la reserva se consideraron para el instrumento del pagare.

Descripción del producto según la Nota Técnica:

- La Prima Anual se calculará:

$$PA = \frac{PU}{\ddot{a}_{x:w-x}} \quad ,$$

dónde:

PU= Prima Única = $A'_{x:w-x}$

- Perdida Esperada del primer año

PE_1 = $Cadq_{NT}$ - PA*α;

dónde α corresponde al recargo por concepto de gastos de adquisición.

- Prima de Ahorro que es la diferencia entre la prima neta nivelada y el costo de siniestralidad esperado del primer año:

$PAh_1 = PN_1 - CS_1$;

con $PN = \frac{PA}{FP}$ y $CS = SA * \frac{q_x}{1+i}$,

dónde FP es la Forma de Pago

- Perdida amortizable

$$PA_1 = mín(PE_1, PAh_1)$$

- Anualidad de amortización en cada año de vigencia del plan:

$$AM_t = \frac{(PA_1) * F_x * \ddot{a}_{x+t:\overline{m-t}|}}{\ddot{a}_{x+1:\overline{m-1}|}};$$

donde $F_x = \frac{(1+i)}{p_x}$ y “m” indica el plazo de pago de primas,

- La reserva mínima del primer año es:

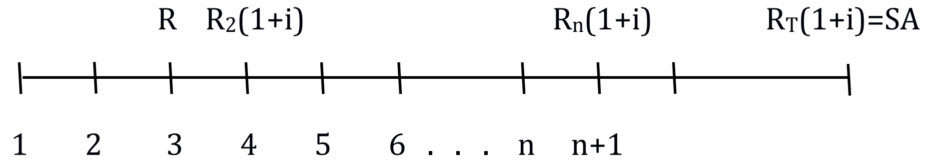
$${}_1V_x^{min} = \frac{\frac{q_x}{1+i} * \frac{365-T}{365} + (PAh_1 - PA_1)(1+i)^{\frac{T}{365}}}{p_x}$$

- Reserva para los demás años:

$${}_tV_x^{min} = {}_tV_x - AM_t$$

La póliza se caracteriza como un dotal mixto, el cual paga al asegurado la Suma Asegurada, si llega vivo al final de la cobertura del seguro o si muere dentro de la vigencia del seguro. Esta vigencia está limitada por una edad final (60 años). Las edades de aceptación del contrato de seguro están sujetas a personas laboralmente activas, con una edad de al menos 22 años.

También me interesa un producto de inversión libre de riesgo, con un plazo correspondiente a la esperanza de vida. Este lo obtengo con el análisis de supervivencia, ya que son de interés los clientes que sobreviven hasta el final, pues con ellos se aumenta el rendimiento de la reserva establecido en el contrato de seguros. Asimismo los plazos de inversión y pago de primas deben coincidir de tal forma que la inversión del instrumento aumente en cada plazo. Esta inversión se denotara por R_t, es decir el rendimiento generado al tiempo t.



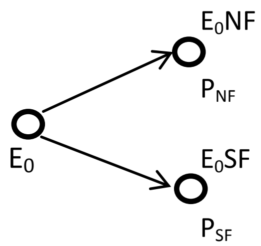
En la línea de tiempo, se puede observar el modelo financiero del seguro, donde la inversión de la reserva se empieza hacer a partir del año 3 de la póliza cuando los gastos son menores. Al tiempo “n” donde se recibe el último pago de prima se debe continuar invirtiendo para que al tiempo T se tenga el monto de la suma asegura, es decir, al final del contrato.

Para conocer el comportamiento necesario que el asegurado debe presentar para sobrevivir hasta el final del contrato de seguro, pues se quiere obtener el resultado mencionado en el modelo financiero, les presentare un enfoque muy utilizado en la actualidad, es decir, Árboles binarios, sin embargo para no caer en lo mismo se utilizará el modelo oculto de Markov o cadena de Markov oculta (también conocida como Red Bayesiana Dinámica [10]).

Árboles Binarios

Definición 1: Un árbol binario es un diagrama que representa diferentes patrones posibles que pueden ser seguidos por las variables durante su tiempo de vida, es decir, los posibles valores que puede tomar la variable mientras transcurre el tiempo. El supuesto más importante en este método es que la variable sigue un camino al azar. En cada paso, se tiene cierta probabilidad de moverse hacia arriba o hacia abajo. En el límite, cuando el tiempo en los pasos es menor, este modelo se basa en el supuesto de que se comporta lognormal[21] que es la razón del modelo de Black-Scholes.

Ahora si en el método los argumentos son no-arbitrarios, para considerar en nuestro caso el hábito de fumar para una persona que supongo empieza su póliza sin fumar E_0, además de que la póliza de la persona dura un tiempo “T” y que en esta duración de tiempo su hábito puede ser que “Si Fumar” pasando al siguiente nivel, E_0SF, donde $SF < 1$, o que “No Fumar”, siendo el siguiente nivel, E_0NF, con $NF > 1$. El valor que toma la variable cuando la persona fuma es 1-SF, disminuyendo el objetivo que queremos alcanzar, pero si el cliente sigue sin fumar entonces se tiene NF-1, que incremente las posibilidades de llegar vivo. Si la persona decide “No Fumar”, suponemos que llegara al siguiente año con una probabilidad favorable P_{NF}, si decide “Si Fumar”, supondremos que llegara al siguiente año con una probabilidad poco favorable P_{SF}.



Como antes, imaginemos un grupo de personas que se mantiene en una posición por cierto tiempo Δ y otras posiciones en una opción. Calculamos el valor de Δ que hace al grupo menos riesgoso. Y hay un movimiento de “No Fumar”, entonces el valore de las personas al final de la póliza es:

$$E_0NF\text{-}P_{NF}$$

Y si el movimiento es "Si Fumar", el valor se vuelve:

$$E_0SF\text{-}P_{SF}$$

Las dos son iguales cuando:

$$E_0NF\text{-}P_{NF} = E_0SF\text{-}P_{SF}$$

ó $$\Delta=\frac{P_{NF}-P_{SF}}{E_0NF-E_0SF}$$

En este caso, es menor el riesgo y, para que no haya oportunidad de arbitrajes se debe ganar al menos el interés acordado[21].

Sin embargo esto sería viable si el Fumar o No fumar tuvieran valores positivos y negativos como en las opciones de inversión, pero como son probabilidades de estados de una persona ambos toman valores positivos por tal motivo seleccione el modelo oculto de Markov. Para poder entender mejor lo que es un modelo oculto de Markov daremos como introducción lo que es una cadena de Markov, y saber cómo se llegó a las cadenas de Markov ocultas.

Cadenas de Markov

Definición 2: Una cadena de Markov $\boldsymbol{q}=\{\boldsymbol{q}_t\}_{t\in N}$ es un proceso estocástico de Markov discreto. Un proceso estocástico se llama de Markov si conocido el presente, el futuro no depende del pasado, esto quiere decir, que dada una variable estocástica $\boldsymbol{q}_{t-1}$ que denota el estado del proceso en el tiempo *t-1*, entonces la probabilidad de transición en el momento *t* se define como $P[q_t = \sigma_t | q_{t-1} = \sigma_{t-1}]$. Formalmente, una cadena de Markov se define como (***Q***,***A***), donde ***Q***={1,2,…,N} son los posibles estados de la cadena y ***A***=$(a_{ij})_{n*n}$es una matriz de transición de estados en el modelo. Si ***A***(t) =$a_{ij}(t)_{n*n}$es independiente del tiempo entonces el proceso se llama homogéneo y las probabilidades de transición de estados son de la forma: $a_{ij}(t) = P[q_t = j | q_{t-1} = i]$ con las siguientes propiedades;

i) $0 \le a_{ij} \le 1$, $1 \le i,j \le N$,

ii) $\sum_{j=1}^{N} a_{ij} = 1$, $1 \le i \le N$.

La condición fundamental de que sea una cadena de Markov establece que las probabilidades de transición y emisión dependen solamente del estado actual y no del pasado, esto es, $P[q_t = j | q_{t-1} = i, q_{t-2} = k, \dots] = P[q_t = j | q_{t-1} = i] = a_{ij}(t)$.

En el presente trabajo considero que el conjunto de estados es finito.

Modelos Ocultos de Markov

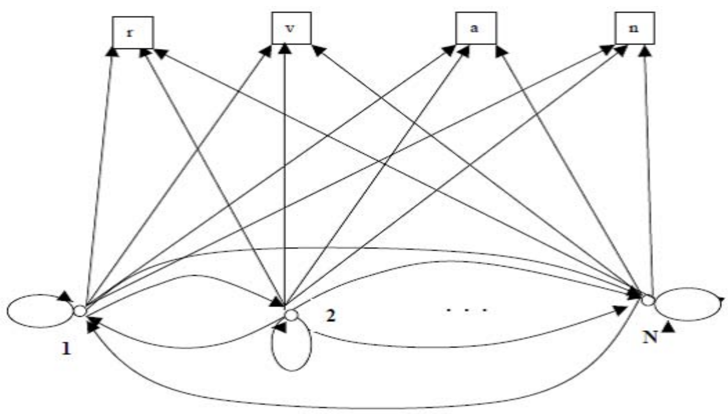
Definición 3: Un Modelo Oculto de Markov (Hidden Markov Model HMM) es un proceso estocástico que consta de un proceso de Markov no observado (oculto) $\boldsymbol{q}=\{\boldsymbol{q}_t\}_{t\epsilon N}$ y un proceso observado $\boldsymbol{O}=\{\boldsymbol{o}_t\}_{t\epsilon N}$ cuyos estados son dependientes estocásticamente de los estados ocultos, es decir, es un proceso bivariado ($\boldsymbol{q}$,$\boldsymbol{O}$). Los HMMs se pueden considerar también como sistemas generativos estocásticos, los cuales se emplean en la modelación de series de tiempo. Un HMM es el tipo más simple de Red Bayesiana Dinámica (RBD) con el que uno se puede encontrar. Se puede definir como un autómata finito estocástico, donde cada estado de proceso está determinado por una única variable aleatoria discreta, donde los valores posibles de las variables son los estados posibles del mundo [10]. Un modelo oculto de Markov es una cadena de $\boldsymbol{q}$ junto con un proceso estocástico que toma valores en un alfabeto **Σ** y el cual depende de $\boldsymbol{q}$.

Estos sistemas evolucionan en el tiempo pasando aleatoriamente de estado a estado y emitiendo en cada momento al azar algún símbolo del alfabeto **Σ**. Cuando se encuentra en el estado $q_{t-1} = i$, tiene la probabilidad a_{ij} de moverse al estado $q_t = j$ en el siguiente instante y la probabilidad $b_j(k)$ de emitir el símbolo $o_t = v_k$ en el tiempo *t.*

Solamente los símbolos emitidos por el proceso $\boldsymbol{q}$ son observables, pero no la ruta o secuencia de estados $\boldsymbol{q}$, de ahí el calificativo de "oculto" de Markov, ya que el proceso de Markov $\boldsymbol{q}$ es no observado. Dando pie al tema que nos interesa para aplicar en el sistema de administración de riesgos actuariales.

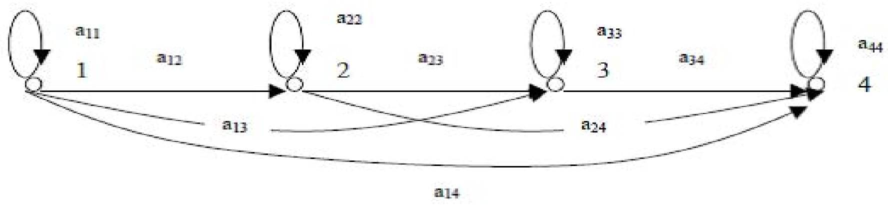
Arquitecturas de los Modelos Ocultos de Markov

Un Modelo Oculto de Markov puede ser representado como un grafo dirigido de transiciones/emisiones como se ilustra en la figura siguiente figura:



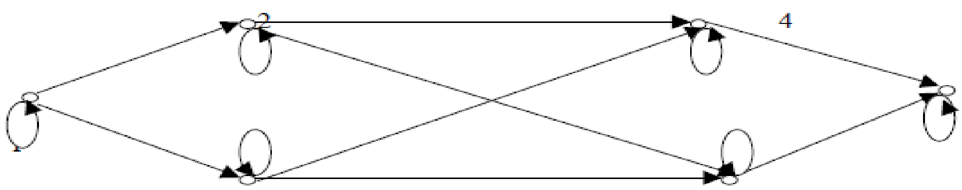
La arquitectura que permita modelar de la mejor forma posible las propiedades observadas depende en gran medida de las características del problema. Las arquitecturas más usadas son:

1) *Ergódicas o completamente conectadas* en las cuales cada estado del modelo puede ser alcanzado desde cualquier otro estado en un número finito de pasos.
2) *Izquierda-derecha, hacia delante o Bakis* las cuales tienen la propiedad de que en la medida que el tiempo crece se avanza en la secuencia de observación asociada ***O,*** y en esa misma medida el índice que señala el estado del modelo permanece o crece, es decir, los estados del sistema van de izquierda a derecha.



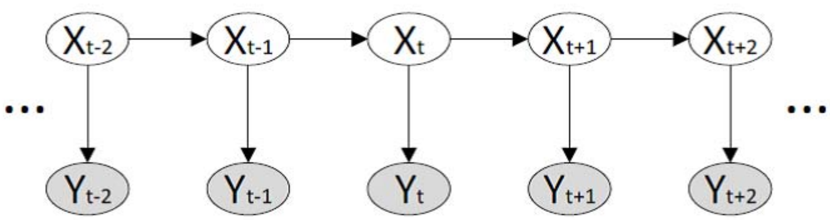
Modelo izquierda-derecha con 4 estados

3) *Izquierda-derecha paralelas,* son dos arquitecturas i*zquierda-derecha* conectadas entre sí.



Modelo izquierda derecha paralelo con 6 estados

Por ejemplo, una red bayesiana correspondiente a un HMM de es una red bayesiana de primer orden (el estado actual depende sólo del estado previo y no de estados iniciales), representada mediante un grafo acíclico dirigido, donde los nodos representan las variables aleatorias (los estados ocultos sin estar sombreados y las observaciones sombreadas) y donde la ausencia de una flecha entre dos variables nos indican su independencia condicional, se ve en la siguiente figura:



Ejemplo de un HMM de primer orden en forma de red bayesiana.

Definición de los elementos de un HMM

Formalmente, un HMM discreto de primer orden se define como una cinco-tupla

$$\boldsymbol{\lambda}=(\boldsymbol{\Sigma},\boldsymbol{Q},\boldsymbol{A},\boldsymbol{B},\boldsymbol{\pi})$$

donde:

i) **Σ**=$\{v_1, v_2, \dots, v_M\}$ es un alfabeto o conjunto discreto finito de *M* símbolos.

ii) ***Q***=*{1,2,…,N}* es un conjunto finito de *N* estados.

iii) ***A***= $(a_{ij})_{N*N}$ es una matriz de probabilidades de transición donde a_{ij} es la probabilidad de transición desde el estado *i* al estado *j*, para todo *i,j* ϵ *N*.

iv) ***B***=($b_j(o_t))_{N*M}$es un vector de probabilidades de emisión de símbolos, uno por cada estado, donde $b_j = (b_{j1}, b_{j2,} \dots, b_{jM})$ es la probabilidad de emisión del símbolo v_k del alfabeto en el estado *j*.

v) ***π***=$(\pi_1, \pi_2, \dots, \pi_N)$ es un vector de probabilidades del estado inicial q_0 en ***Q****.*

Las probabilidades de iniciación, transición y emisión son los parámetros del modelo.

Un HMM poder ser usado como un generador de secuencias de observaciones ***O***=$(o_1, o_2, \dots, o_T)$ donde:

- $o_t$es uno de los símbolos de **Σ** para *tϵ{1,2,…,T}*
- Tes la longitud de la secuencia de observaciones ***O***, es decir, el número de observaciones en la secuencia.
- **λ**=(***A***,***B***,***π***) son los parámetro s del modelo.

Un HMM define una medida de probabilidad *μ*, sobre el espacio de secuencias ***Σ****.

Problemas básicos de los HMMs

Existen tres problemas básicos relacionados con los HMMs:

1. Calcular eficientemente $P(\boldsymbol{O}|\lambda)$ la probabilidad de la secuencia de observación ***O*** dado el modelo **λ**=(***A***,***B***,***π***) y la secuencia de observación ***O***=$(o_1, o_2, \dots, o_T)$.

2. Encontrar la trayectoria más probable $q=(q_1q_2...q_T)$ dado el modelo λ y la secuencia de observación $\boldsymbol{O}=(o_1, o_2, \dots, o_T)$, es decir, $q=arg_{r\epsilon Q^*}\{max_{r\epsilon Q^*}P(r)\}$.

3. Ajustar los parámetros $\boldsymbol{A}$,$\boldsymbol{B}$,$\boldsymbol{\pi}$ para maximizar $P(\boldsymbol{O}|\lambda)$.

Estimación inicial de los parámetros del HMM

Existen diferentes formas de incorporar la información inicial para el diseño del HMM y estimar sus parámetros. La experiencia ha mostrado que las estimaciones iniciales uniformes o aleatorias de los parámetros **π** y ***A*** son adecuados en la mayoría de los casos y que las buenas estimaciones iniciales para el parámetro ***B*** son útiles en el caso de secuencias con símbolos discretos y esenciales en el caso de distribuciones continuas.

Evaluación de las probabilidades de la secuencia [10]

El Problema 1 de los HMMs consiste en calcular la probabilidad de la secuencia de observación $\boldsymbol{O}=(o_1, o_2, \dots, o_T)$, dado el modelo $\boldsymbol{\lambda}$, es decir, $P(\boldsymbol{O}|\lambda)$. La forma más simple de resolver el Problema 1 consiste en enumerar todas las posibles secuencias de estado de longitud ***T***. Una de dichas secuencias es de la forma $q=(q_1q_2...q_T)$ donde q_1 es el estado inicial.

La probabilidad de la secuencia de observación ***O*** dada la anterior secuencia ***q*** es:

$$\boldsymbol{P(O|q_0\lambda)} = \prod_{t=1}^{T} P(o_t|q_t,\lambda) = b_{q_1}(o_1)b_{q_2}(o_2)\dots b_{q_T}(o_T)$$

La probabilidad de la secuencia de estados ***O*** es:

$$\boldsymbol{P(q|\lambda)} = \boldsymbol{\pi}_{q1}a_{q1\,q2}a_{q2\,q3}\dots a_{qT-1\,qT}$$

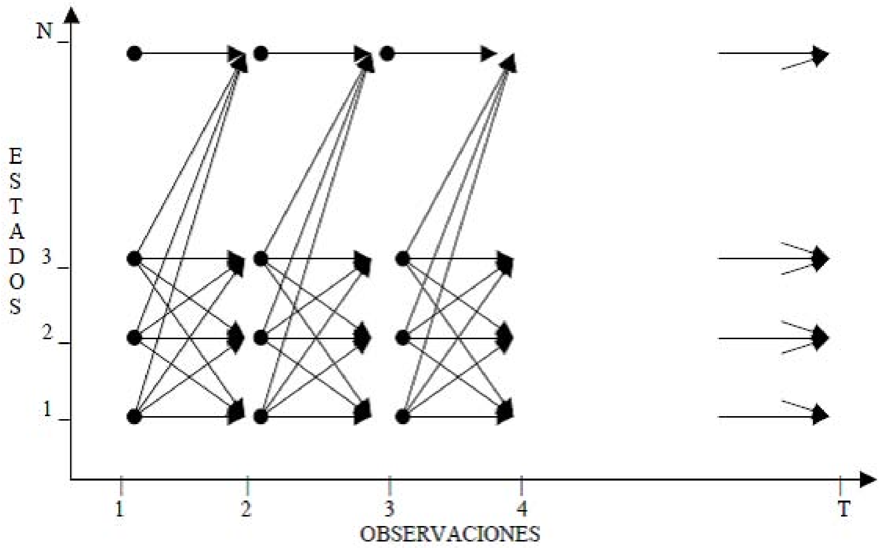
La probabilidad de que ***O*** y *q* ocurran simultáneamente está dada por:

$$P(\boldsymbol{O},\boldsymbol{q}|\boldsymbol{\lambda}) = P(\boldsymbol{O}|\boldsymbol{q},\boldsymbol{\lambda})P(\boldsymbol{q}|\boldsymbol{\lambda})$$

La probabilidad de ***O*** sobre todas las posibles secuencias de estados ***Q***, es el cálculo de:

$$P(\boldsymbol{O}|\boldsymbol{\lambda}) = P(\boldsymbol{O}|\boldsymbol{q},\boldsymbol{\lambda})P(\boldsymbol{q}|\boldsymbol{\lambda}) = \boldsymbol{\pi}_{q1}b_{q1}(o_1)a_{q1\,q2}b_{q2}(o_2)\dots a_{qT-1\,qT}b_{qT}(o_T)$$

Esta expresión nos permite un cómputo eficiente de la probabilidad debido a que el número de rutas en una arquitectura es exponencial. En cada tiempo t=1,2,…,$\boldsymbol{T}$ se tienen $\boldsymbol{N}$ posibles estados actuales, por lo tanto $\boldsymbol{N}^{\boldsymbol{T}}$ operaciones, como se puede observar en la siguiente figura.



Existe un procedimiento más eficiente para calcular dicha probabilidad, denominado *algoritmo de avance (forward algorithm).*

Proposición 1: Algoritmo de avance (forward algorithm)[10]

Sea $\alpha_t(i) = P(o_1 o_2 \dots o_t, q_t = i|\boldsymbol{\lambda})$

La probabilidad de la secuencia de observación parcial $o_1 o_2 \dots o_t$ en el estado *i* hasta el tiempo t, dado el modelo $\boldsymbol{\lambda}$. Puede calcularse $\alpha_t(i)$, así:

1. *Inicializa*: $\alpha_1(i) = \pi_i b_i(o_1)$ $1 \le i \le \boldsymbol{N}$,
2. Inducción $\alpha_{t+1}(j) = \sum_{i=1}^{N} \alpha_t(i) a_{ij} b_j(o_{t+1})$ $1 \le t \le \boldsymbol{T}$-1, $1 \le j \le \boldsymbol{N}$,

Demostración 1: $\alpha_t(i) = P(o_1 o_2 \dots o_t, q_t = i|\boldsymbol{\lambda})$, haciento t = 1 tenemos

$$\alpha_1(i) = P(o_1, q_1 = i|\lambda) = P(o_1, q_1 = i)P(q_1 = i) = b_i(o_1)\pi_i \blacksquare$$

Demostración 2: $\alpha_{t+1}(j) = P(o_1 o_2 \dots o_{t+1}, q_{t+1} = j|\boldsymbol{\lambda})$

$$= \sum_{i=1}^{N} P(o_1 o_2 \dots o_{t+1}, q_t = i, q_{t+1} = j|\boldsymbol{\lambda})$$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_t, q_t = i|\boldsymbol{\lambda}) P(o_{t+1}, q_{t+1} = j|o_1 o_2 \dots o_t, q_t = i, \boldsymbol{\lambda})$

$=\sum_{i=1}^{N} \alpha_t(i) P(o_{t+1}, q_{t+1} = j|q_t = i, \boldsymbol{\lambda})$

$=\sum_{i=1}^{N} \alpha_t(i) P(o_{t+1}|q_{t+1} = j, q_t = i, \boldsymbol{\lambda})\, P(q_{t+1} = j|q_t = i, \boldsymbol{\lambda})$

$=\sum_{i=1}^{N} \alpha_t(i) P(o_{t+1}|q_{t+1} = j, \boldsymbol{\lambda})\, a_{ij}$

$=\sum_{i=1}^{N} \alpha_t(i) b_j(o_{t+1})\, a_{ij}$ ∎

3. *Terminación*: $P(\boldsymbol{O}|\boldsymbol{\lambda}) = \sum_{i=1}^{N} \alpha_T(i)$

Demostración: $\sum_{i=1}^{N} \alpha_T(i) = \sum_{i=1}^{N} P(o_1 o_2 \dots o_T, q_T = i|\boldsymbol{\lambda})$

$= \sum_{i=1}^{N} P(o_1 o_2 \dots o_T, q_{T-1} = i, q_T = i|\boldsymbol{\lambda})$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-1}, q_{T-1} = i|\boldsymbol{\lambda}) P(o_T, q_t = i|o_1 o_2 \dots o_{T-1}, q_{T-1} = i, \boldsymbol{\lambda})$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-1}, q_{T-1} = i|\boldsymbol{\lambda}) P(o_T, q_T = i|q_{T-1} = i, \boldsymbol{\lambda})$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-1}, q_{T-1} = i|\boldsymbol{\lambda}) P(o_T|q_T = i, q_{T-1} = i, \boldsymbol{\lambda}) P(q_T = i|q_{T-1} = i, \boldsymbol{\lambda})$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-1}, q_{T-1} = i|\boldsymbol{\lambda}) P(o_T|q_T = i, \boldsymbol{\lambda}) a_{q_T q_{T-1}}$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-1}, q_{T-2}, q_{T-1} = i|\boldsymbol{\lambda}) b_{q_T}(o_T) a_{q_T q_{T-1}}$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-2}, q_{T-2} = i|\boldsymbol{\lambda}) P(o_{T-1}, q_{T-1} = i|q_{T-2} = i, \boldsymbol{\lambda}) b_{q_T}(o_T) a_{q_T q_{T-1}}$

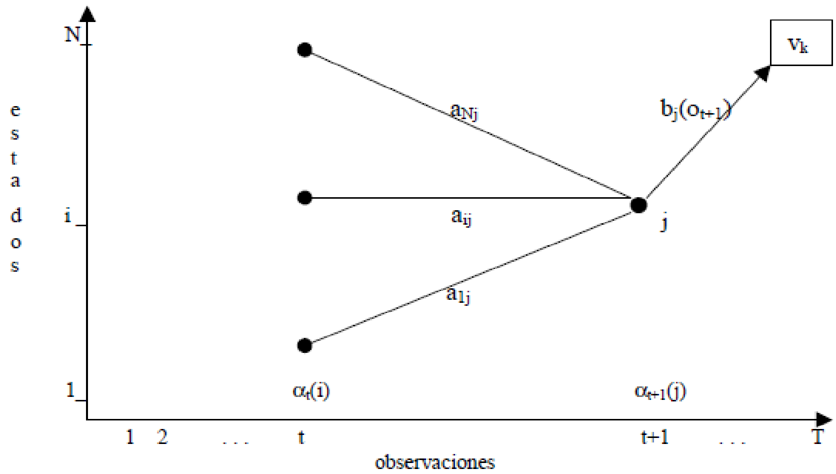
$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-2}, q_{T-2} = i|\lambda) P(o_{T-1}|q_{T-1} = i, \lambda) P(q_{T-1} = i|q_{T-2} = i, \lambda) b_{q_T}(o_T) a_{q_T q_{T-1}}$

$=\sum_{i=1}^{N} P(o_1 o_2 \dots o_{T-2}, q_{T-2} = i|\lambda) b_{q_{T-1}}(o_{T-1}) a_{q_{T-1} q_{T-2}} b_{q_T}(o_T) a_{q_T q_{T-1}}$

$=\dots= P(o_1, q_1 = i|\boldsymbol{\lambda}) b_{q_2}(o_2) a_{q_2 q_1} \dots b_{q_T}(o_T) a_{q_T q_{T-1}} = \pi_{q_1} b_{q_1}(o_1) b_{q_2}(o_2) a_{q_2 q_1} \dots b_{q_T}(o_T) a_{q_T q_{T-1}}$

$= P(\boldsymbol{O}|\boldsymbol{\lambda})$ ∎

El número de operaciones requeridas para calcular $\alpha_t(j)$, $1 \leq t \leq \boldsymbol{T}$, $1 \leq j \leq \boldsymbol{N}$, ilustradas en la siguiente figura, que son exactamente igual a ***N(N+1)(T-1)+N*** multiplicaciones y ***N(N-1)(T-1)*** sumas. Entonces es del orden $\boldsymbol{N^2 T}$.



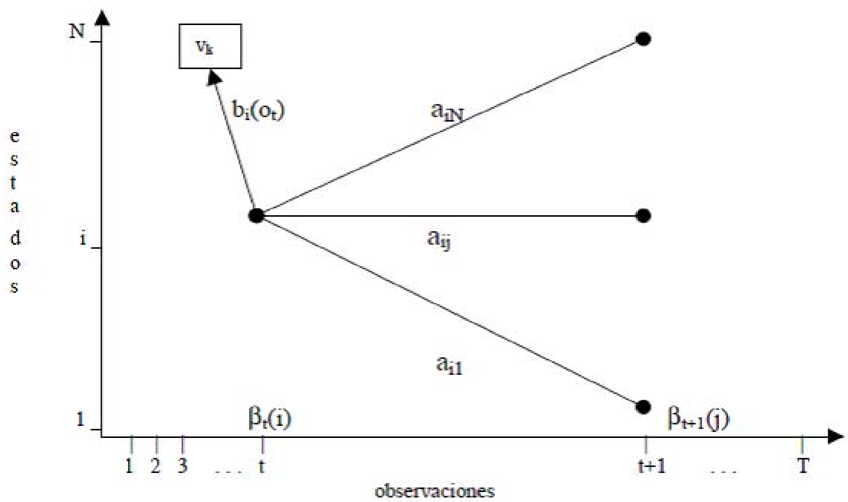
Proposición 2: Algoritmo de retroceso (backward algorithm)[10]

Sea $\beta_t(i) = P(o_{t+1}o_{t+2} \dots o_T | q_t = i, \boldsymbol{\lambda})$

La probabilidad de la secuencia de observación parcial desde *t+1* hasta el final, dado el estado *i* en el tiempo *t* y el modelo $\boldsymbol{\lambda}$. La variable $\beta_t(i)$ puede resolver como sigue:

1. *Inicializa* $\beta_t(i) = 1, \quad 1 \leq i \leq \boldsymbol{N},$
2. *Inducción:* $\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(o_{t+1}) \beta_{t+1}(j), \qquad t = T-1, T-2, \dots, 1 \; y \; 1 \leq i \leq \boldsymbol{N}$

Los cómputos necesarios de $\beta_t(i) = 1, 1 \leq t \leq \boldsymbol{T}, 1 \leq i \leq \boldsymbol{N},$ es del orden $\boldsymbol{N^2 T}$ operaciones, según se ilustra en la siguiente figura:



Secuencia optimal de estados

Definición 4 (Algoritmo de Viterbi): La ruta más probable *q= ($q_1q_2…q_T$)* para la secuencia de observación ***O***= (*$o_1o_2…o_T$)* puede ser calculada utilizando el *Algoritmo de Viterbi*, la cual es útil para el aprendizaje y para el alineamiento de secuencias en el modelo de un HMM.

$$\text{Sea } \delta_t(i) = \max_{q_1,q_2,\dots,q_{t-1}} P(q_1q_2 \dots q_{t-1}, q_t = i, o1o2 \dots ot|\boldsymbol{\lambda})$$

la probabilidad asociada con la ruta más probable a lo largo de un camino simple que toma en cuenta las primeras *t* observaciones y finaliza en el estado *i*. Por inducción se tiene que

$$\delta_{t+1}(j) = \max_i \delta_t(i) a_{ij} b_j(o_{t+1})$$

Para conservar la secuencia de estados se define la variable $\Psi_t(j)$. El procedimiento para encontrar la mejor secuencia de estados se establece de la siguiente manera:

1. *Inicialización:* $\delta_1(i) = \pi_i b_i(o_1)$ $1 \leq i \leq \boldsymbol{N}$

   $\Psi_1(i) = 0,$ $1 \leq i \leq \boldsymbol{N}$

2. *Inducción:* $\delta_t(j) = \max_{1\leq i\leq N} \delta_{t-1}(i) a_{ij} b_j(o_t)$ $2 \leq t \leq \boldsymbol{T}, 1 \leq j \leq \boldsymbol{N}$

   $\Psi_t(j) = arg \max_{1\leq i\leq N} \delta_{t-1}(i) a_{ij},$ $2 \leq t \leq \boldsymbol{T}, 1 \leq j \leq \boldsymbol{N}$

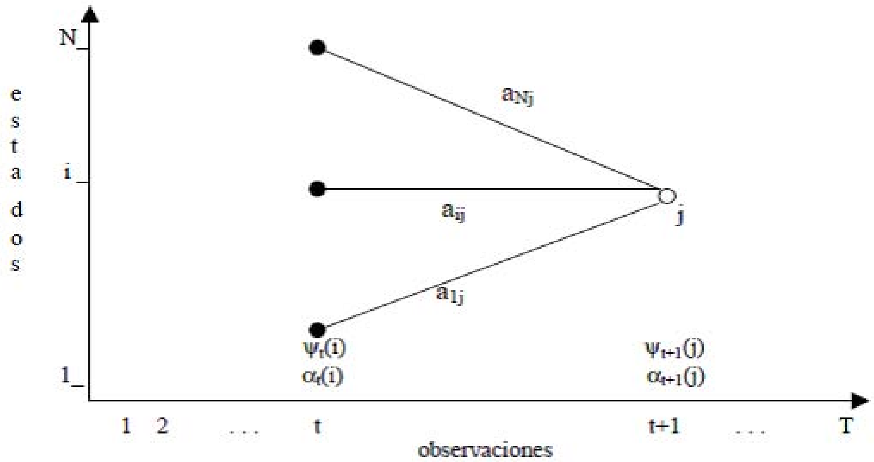
3. *Terminación:* $P^* = \max_{1\leq i\leq N} \delta_T(i)$ *y* $Q_T^* = arg \max_{1\leq i\leq N} \delta_T(i)$

4. *Ruta inversa:* $q_t^* = \Psi_{t+1}(q_{t+1}^*)$, *t*=*T*-1,*T*-2,…,1.

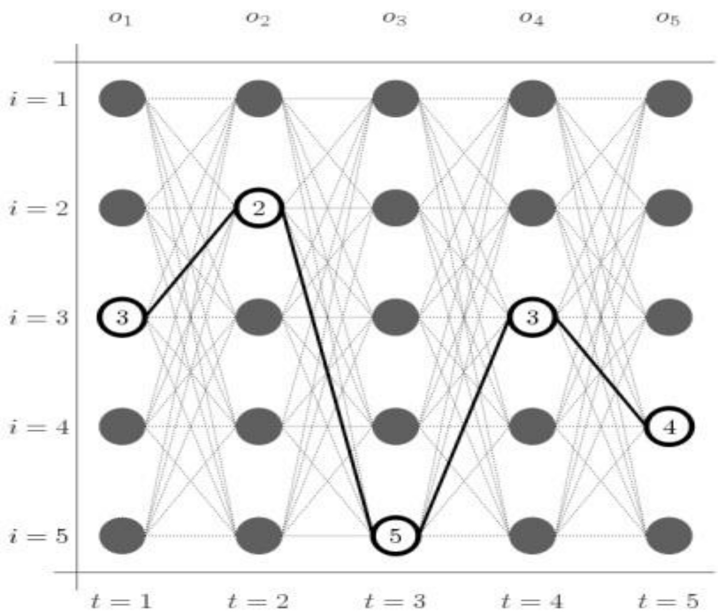
El algoritmo de Viterbi permite encontrar las secuencias de estados más probables en un Modelo Oculto de Markov, es decir, obtiene la secuencia óptima que mejor explica las observaciones, donde $\delta_t(j)$ es la probabilidad del mejor camino hasta el estado i habiendo visto las t primeras observaciones, esta función se calcula para todos los estados e instantes de tiempo.

Dado que el objetivo es obtener la secuencia de estados más probables, se necesita almacenar el argumento que maximice el valor de la ecuación en cada instante de tiempo t y para cada estado j, por lo que el algoritmo de Viterbi se basa en el método de la programación dinámica.

La estructura de retículo implementa eficientemente los cómputos del algoritmo, como se puede ver en la siguiente figura:



Como ejemplo se encuentra la siguiente figura, con la secuencia de estados más probables en un Modelo Oculto de Markov de 5 estados dada una secuencia de observaciones de longitud 5.



Aprendizaje del modelo

El problema más difícil de los HMMs es determinar un método para ajustar los parámetros (***A, B, π***) del modelo para satisfacer los criterios de optimización. No se conoce una forma analítica para fijar los parámetros que maximice la probabilidad de la secuencia de observación, sin embargo existen distintos algoritmos para el entrenamiento de un HMM, entre ellos, Baum-Welch o EM (Expectation Maximization).

El algoritmo de Baum-Welch[11]

Definición 5: Intuitivamente el algoritmo de Baum-Welch podría describirse de la siguiente manera. En un comienzo no se conocen los parámetros que mejor ajustan al modelo por lo que se eligen en forma aleatoria, pero se dispone de una o más secuencias de entrenamiento *x* que pueden ser utilizadas para maximizar la probabilidad de transición y emisión. Con los parámetros escogidos en el paso de inicio, se identifican las transiciones y emisiones más probables para el conjunto de entrenamiento que suplantan entonces a los valores iniciales y como consecuencia se mejora el modelo, es decir, se obtiene un modelo con mayor probabilidad de haber generado la secuencia dada. El proceso se repite hasta que no se observan mejoras en el modelo.

Si x es una secuencia de entrenamiento, la probabilidad de que la transición del estado k al estado l sea utilizada en la posición i de x es

$$P(y_i = k, y_{i+1} = l|x) = \frac{P(y_i = k, y_{i+1} = l, x)}{P(x)}$$

$$= \frac{P(x_1, \dots, x_i, y_i = k)P(y_{i+1} = l|y_i = k)P(x_{i+1}|y_{i+1} = l)P(x_{i+2}, \dots, x_n|y_{i+1} = l)}{P(x)}$$

$$= \frac{\alpha_k(i)P(y_{i+1} = l|y_i = k)P(x_{i+1}|y_{i+1} = l)\beta_l(i+1)}{P(x)}$$

Dónde $\alpha_k(i) = P(x_1, \dots, x_i, y_i = k)$ y $\beta_l(i+1) = P(x_{i+2}, \dots, x_n|y_{i+1} = l)$ son las probabilidades forward y backward definidas anteriormente y la probabilidad del denominador se calcula mediante el algoritmo forward.

Luego, para calcular la proporción esperada de veces que se utiliza la transición del estado k al estado l en la secuencia x, basta con sumar en todos los lugares i

$$\sum_i P(y_i = k, y_{i+1} = l|x) = \sum_i \frac{\alpha_k(i)P(y_{i+1} = l|y_i = k)P(x_{i+1}|y_{i+1} = l)\beta_l(i+1)}{P(x)}$$

De manera análoga se calcula la proporción de veces que el símbolo b es generado por el estado k:

$$\frac{\sum_{\{i:x_i=b\}} \alpha_k(i)\beta_k(i)}{P(x)}$$

Supongamos se cuenta con un número T de secuencias de entrenamiento que llamamos

$$x^t \text{con } t = 1, \dots, T.$$

Por un lado las probabilidades de transición se estiman como el cociente entre el número esperado de transiciones del estado k al l y el número esperado de transiciones desde k en todas las secuencias x^t.Por otra parte las probabilidades de emisión son calculadas como el cociente entre el número esperado de veces que se transita por el estado k y se emite el símbolo b y el número esperado de veces que se pasa por k también en todas las secuencias de entrenamiento.

Luego los estimadores de los parámetros del modelo $\hat{p}_{kl}\ y\ \hat{e}_k(b)$ son:

$$\hat{p}_{kl} = \frac{P_{kl}}{\sum_{l'} P_{kl'}} \quad \text{y} \quad \hat{e}_k(b) = \frac{E_k(b)}{\sum_{b'} E_k(b')}$$

dónde:

$$P_{kl} = \sum_t \frac{1}{P(x^t)} \sum_i \alpha_k^t(i) P(y_{i+1} = l | y_i = k) P(x_{i+1}^t | y_{i+1} = l) \beta_l^t(i+1),\ y$$

$$E_k(b) = \sum_t \frac{1}{P(x^t)} \sum_{\{i: x_i^s = b\}} \alpha_k^t(i) \beta_k^t(i)$$

Ahora, si se suplantan los valores de los parámetros $\hat{\theta}_k$ ($\hat{p}_{kl}\ y\ \hat{e}_k(b)$) por los calculados en el paso anterior entonces, puede probarse que $P_{\hat{\theta}_{h+1}}(x) \geq P_{\hat{\theta}_h}(x)$, es decir, la verosimilitud del modelo aumenta con las iteraciones sucesivas.

El algoritmo llega a su fin o bien cuando la diferencia en las probabilidades de un paso a otro es imperceptible o cuando se alcanza el número de iteraciones fijado (criterio de parada).

El método de Baum-Welch presenta una fuerte dependencia de la elección inicial de los parámetros lo que implica que el algoritmo puede no encontrar el mejor modelo si la inicialización de las probabilidades de transición y emisión no resultan ser las adecuadas.

Resumen del Algoritmo de Baum-Welch

1. *Inicialización:* Se elige un valor arbitrario para θ
2. *Iteración*
   a) cálculo de $\alpha_k^t(i)$, $\beta_k^t(i)$ para cada secuencia *t*, P_{kl} y $E_k(b)$

b) cálculo de $\hat{p}_{kl}$ y $\hat{e}_k(b)$, es decir cálculo de $\hat{\theta}$ (nuevo modelo)

c) cálculo de $P_{\hat{\theta}}(x)$

d) si $P_{\hat{\theta}}(x) > P_{\theta}(x)$ se considera θ = $\hat{\theta}$ y se vuelve al paso a

3. *Finalización*

Cuando $P_{\hat{\theta}}(x) = P_{\theta}(x)$ o se alcanza el número de iteraciones pautado

El algoritmo de Baum-Welch comienza asignando valores iniciales a los parámetros, esto es a las probabilidades de transición y de emisión, que pueden o bien sortearse de una distribución uniforme, o elegirse si se cuenta con algún conocimiento a priori que lo permita. En cada iteración se calcula el número esperado de veces en que cada transición y emisión es utilizada por el modelo para generar los datos de entrenamiento con los parámetros del paso anterior. Luego se actualizan las probabilidades de transición y emisión de forma tal que se maximice la verosimilitud de los valores esperados. El algoritmo EM permite encontrar estimadores de máxima verosimilitud de los parámetros en modelos probabilísticos, aun cuando en el conjunto de entrenamiento se tienen datos faltantes (en el caso de un HMM los datos faltantes serían los estados que se encuentran ocultos).

El proceso iterativo en el algoritmo Baum-Welch se basa en el aumento de la verosimilitud en cada paso, a continuación se mostrará por qué la función de verosimilitud es creciente con el número de iteraciones.

Teorema 1. Desigualdad de Gibbs

Si P_1 es la distribución de probabilidad de una variable aleatoria "*x*" y P_2 es otra distribución de probabilidad discreta entonces

$$d\big(P_1(x), P_2(x)\big) = \sum_i P_1(x_i)\log\left[\frac{P_1(x_i)}{P_2(x_i)}\right] \geq 0$$

Demostración:

Si φ es una función convexa, la desigualdad de Jensen establece que E(φ(y)) ≥ φ(E(y)).
Dado que la función convexa φ(y)=-log(y) es convexa, aplicando la desigualdad de Jensen a la variable aleatoria $y = \frac{P_1(x)}{P_2(x)}$ se tiene que

$$\sum_i P_1(x_i)\log\left[\frac{P_1(x_i)}{P_2(x_i)}\right] \geq -\log\sum_i P_1(x_i)\frac{P_2(x_i)}{P_1(x_i)} = -\log\sum_i P_2(x_i) = 0 \blacksquare$$

Teorema 2(la función Q)

Sean "*x*" e "*y*" las variables aleatorias y la función Q(θ , θ')=$\sum_y P_\theta(y|x)\log P_{\theta'}(x,y)$, donde θ y $\hat{\theta}$ representan los parámetros de dos modelos distintos. Entonces

$$Q(\theta , \theta') \geq Q(\theta , \theta) \;\rightarrow\; P_{\theta'}(x) \geq P_\theta(x)$$

Demostración:

Como $P_\theta(x) = \frac{P_\theta(x,y)}{P_\theta(y|x)}$, tomando logaritmos se tiene que

$$\log P_\theta(x) = \log\left[\frac{P_\theta(x,y)}{P_\theta(y|x)}\right] = \log P_\theta(x,y) - \log P_\theta(y|x)$$

Multiplicando por $P_\theta(y|x)$ y sumando en "*y*":

$$\log P_\theta(x) = \sum_y P_\theta(y|x)\log P_\theta(x,y) - \sum_y P_\theta(y|x)\log P_\theta(y|x)$$

$$= Q(\theta,\theta) - \sum_y P_\theta(y|x)\log P_\theta(y|x)$$

Luego

$$\log P_{\theta'}(x) - \log P_\theta(x)$$
$$= Q(\theta,\theta') - \sum_y P_\theta(y|x)\log P_{\theta'}(y|x) - Q(\theta,\theta) + \sum_y P_\theta(y|x)\log P_\theta(y|x)$$

$$= Q(\theta,\theta') - Q(\theta,\theta) + \sum_y P_\theta(y|x)\log\frac{P_\theta(y|x)}{P_{\theta'}(y|x)}$$

El último sumando resulta no negativo debido a la desigualdad de Gibbs por lo que $\log P_{\theta'}(x) - \log P_\theta(x) \geq Q(\theta,\theta') - Q(\theta,\theta)$, de donde se concluye que si $Q(\theta,\theta') \geq Q(\theta,\theta)$ entonces $\log P_{\theta'}(x) \geq \log P_\theta(x)$, es decir $P_{\theta'}(x) \geq P_\theta(x)$ ∎

El algoritmo de Baum-Welch se fundamenta en el teorema anterior, puesto que, si se encuentra un valor del parámetro θ' para el cual se satisface la primera desigualdad en el Teorema, los datos de entrenamiento "*x*" serán más probables bajo el modelo de parámetro θ' que bajo el modelo de parámetro θ, es decir aumenta la verosimilitud.

De lo anterior, el procedimiento de entrenamiento con EM en un HMM, para reestimar los parámetros ***A***, ***B*** y ***π***, utiliza la variable $\xi_t(i,j)$, que es la probabilidad de encontrarse en el estado *i* en el tiempo *t* y en el estado *j* en el tiempo *t+1*, dado el modelo ***λ*** y la secuencia de observación ***O***, esto es

$$\xi_t(i,j) = P[q_t = i, q_{t+1} = j|\boldsymbol{O},\boldsymbol{\lambda}] = \frac{P[q_t = i, q_{t+1} = j, \boldsymbol{O}|\boldsymbol{\lambda}]}{P[\boldsymbol{O}|\boldsymbol{\lambda}]}$$

La variable $\xi_t(i,j)$ puede reescribirse a partir de las variables $\alpha_t(j)\ y\ \ \beta_t(i)$, así:

$$\xi_t(i,j) = \frac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}$$

La suma de $\xi_t(i,j)$ sobre *t* puede interpretarse como el número esperado de transiciones desde el estado *i* hasta el estado *j*, es decir,

$\sum_{i=1}^{T-1}\sum_{j=1}^{N}\xi_t(i,j) =$ número esperado de transiciones desde el estado i en ***O***

$\sum_{i=1}^{T-1}\xi_t(i,j) =$ número esperado de transiciones desde el estado *i* hasta el estado *j* en ***O***

Con las anteriores expresiones se pueden enunciar las siguientes fórmulas de reestimación para los parámetros ***A***, ***B*** y ***π***.

$\pi_i' =$ frecuencia esperada en el estado *i* en el tiempo *t*=1, entonces es igual a

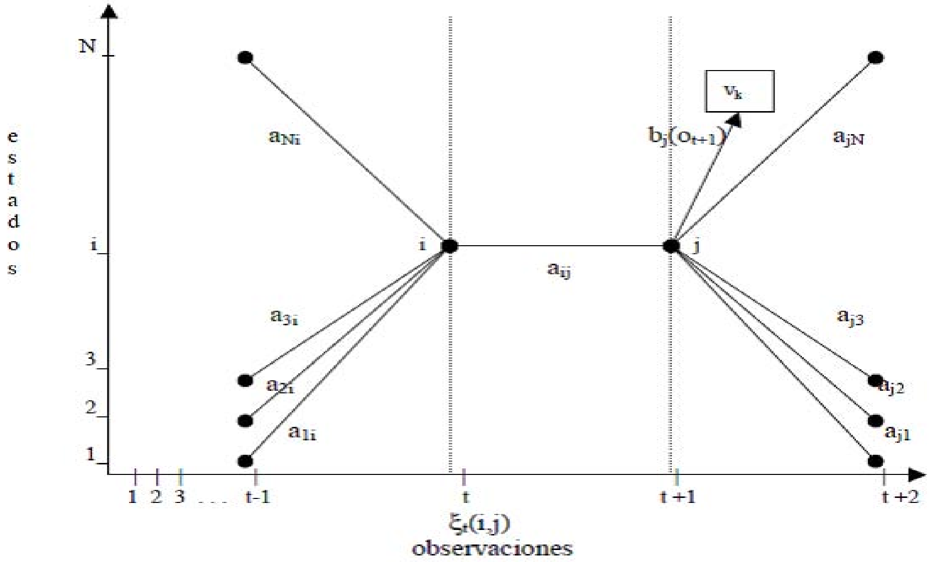
$$\boldsymbol{\pi}_i' = \sum_{i=1}^{N}\xi_1(i,j)$$

$$a_{ij}' = \frac{\text{número esperado de transiciones desde el estado } i \text{ hasta el estado } j}{\text{número esperado de transiciones desde el estado } i}$$

$$a_{ij} = \frac{\sum_{i=1}^{T-1}\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{T-1}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}$$

$$b_j'(k) = \frac{\text{número esperado de veces en el estado } j \text{ y observando el símbolo } v_k}{\text{número esperado de veces en el estado } j}$$
$$= \frac{\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(j)a_{ij}b_j(o_{t+1})\beta_{t+1}(j),\ \ o_t = v_k}{\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}$$

Las trayectorias que satisfacen las condiciones requeridas por la ecuación se ilustran en la siguiente figura:

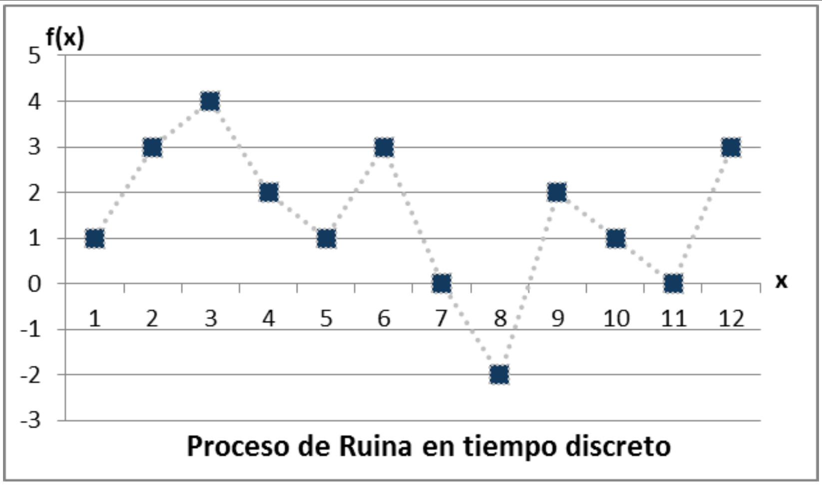


Si designamos al modelo reestimado como ***λ'****= (**A', B'**, **π'**),* pueden presentarse dos posibilidades:

1) El modelo inicial (***A, B, π****)* define un punto crítico de la función de probabilidad, en cuyo caso, ***λ'****=* ***λ****; ó*

2) El modelo ***λ'*** es más probable que el modelo ***λ***, es decir, $P(\boldsymbol{O}|\lambda') > P(\boldsymbol{O}|\lambda)$ y se ha conseguido un nuevo modelo ***λ'*** a partir del cual la secuencia de observación es más probable.

Con respecto al hábito de fumar, es necesario definir una tendencia dinámica sobre dicho hábito y saber cómo afecta a la siniestralidad durante el periodo del contrato de seguro. Es decir, encontrar el comportamiento deseado del asegurado con respecto al hábito de fumar para que éste sobreviva hasta el final del contrato. Para esto, se plantea el uso del modelo de la Cadena Oculta de Markov, que ayuda a definir el comportamiento del hábito de fumar mediante una Cadena de Markov y las observaciones asociadas como los estados del evento de siniestralidad, la cual, será una variable binaria.



Dado que se desea que no suceda la siniestralidad durante el tiempo que dura el contrato de seguros (se puede ver como un caso análogo de no querer llegar a la ruina), para obtener este resultado se aplicara el siguiente procedimiento:

1.- Establecer un modelo inicial de la cadena de Markov Oculta

2.-Establecer la tendencia S1=S2=….=Sn=0. Esta tendencia significa que no se tiene siniestralidad durante el contrato del seguro.

3.- Se ejecuta el algoritmo de Baum Welch para determinar el Modelo Markoviano oculto λ' que maximice la probabilidad P(S1=S2=….=Sn=0|λ').

4.- Se aplica el algoritmo de Viterbi para encontrar la secuencia de estados de la Cadena de Markov que maximiza la probabilidad P(S1=S2=….=Sn=0|λ').

El comportamiento encontrado por el paso 4, representa entonces la tendencia deseada del asegurado al fumar, y con probabilidad P(F1,…., Fn|λ')

IV.- Estimación del modelo al caso de estudio

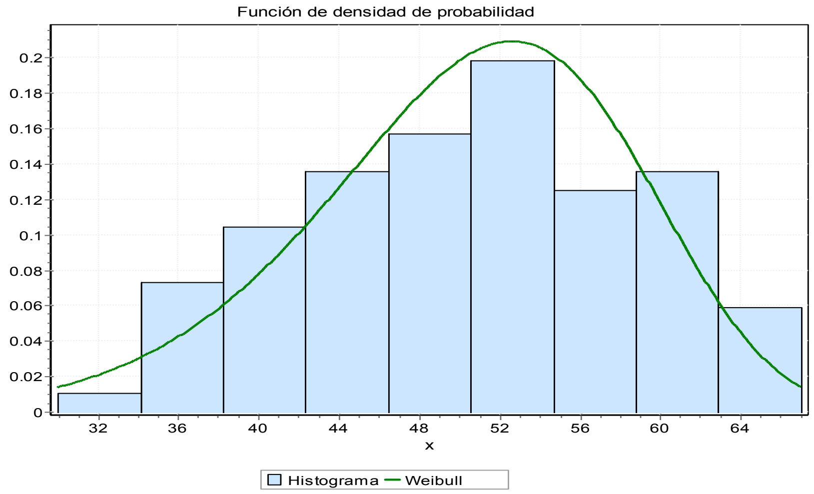
Dado el modelo financiero se asume la no ocurrencia del siniestro mientras esté vigente el contrato de seguro, es decir que ocurra el siniestro después del final del contrato. Dado este supuesto se habrán de encontrar los parámetros que incorporen la información de los factores de riesgo, en este caso, la edad y el fumar. Para el caso de edad se aplica un análisis de supervivencia para saber la edad promedio de supervivencia y para determinar una regla discriminatoria para la edad mínima de contratación.

Posteriormente se aplicara un modelo Markoviano oculto para el factor fumar, es decir, encontrar el comportamiento deseable del asegurado durante el tiempo del plazo de la inversión financiera.

Los datos que se utilizaran en el análisis de la base de siniestros, por políticas de la aseguradora y cubrir los datos del cliente son las siguientes: edad de entrada, fumador, edad de salida, edad final. A continuación se presentan en forma de tabla con su descripción respectiva:

Nombre	Descripción	Símbolo	tipo	tipo R	subtipo	Escala	Espacio de estados
Edad de entrada	Edad del cliente al contratar el seguro	edadi	cuantitativa	Numeric	Discreta	años	[22,50]
Fumador	Tipo de habito del cliente	fuma	cualitativa	String	Nominal	booleano	[S, N]
Edad Final	Edad en que termina el seguro	edadf	cuantitativa	Numeric	Discreta	años	[55,65]
Edad de Salida	Edad en la que el asegurado da fin el contrato de seguro	edads	cuantitativa	Numeric	Discreta	años	[22,50]

Ahora se hace el análisis de la información para saber que distribución siguen las edades de salida y aplicar la función de supervivencia correspondiente. Para ello primero se aplican las pruebas de bondad de ajuste con ayuda del programa EasyFit [20] (Calcula los parámetros por medio del método de máxima verosimilitud y en los casos que no se puede despejar el estimador de la función de verosimilitud utiliza aproximaciones numéricas), el cual arroja la función de distribución que mejor se ajusta a nuestros datos, a continuación se muestra el histograma de salida.



Con el histograma se puede observar que la función de distribución Weibull se ajusta bien a la información de la variable “edades de salida”.

Para corroborar este ajuste presento los resultados de las pruebas de bondad de ajuste que también obtuve del programa Easyfit, en la siguiente tabla:

Bondad de ajuste – Detalles

Weibull

Kolmogorov-Smirnov

Tamaño de la muestra	287			
Estadística	0.06671			
Valor P	0.14856			
Rango	23			
⍰	0.2	0.1	0.05	0.02
Valor crítico	0.06334	0.07219	0.08016	0.0896
Rechazar?	Sí	No	No	No

Anderson-Darling

Tamaño de la muestra	287			
Estadística	1.7759			
Rango	24			
⍰	0.2	0.1	0.05	0.02
Valor crítico	1.3749	1.9286	2.5018	3.2892
Rechazar?	Sí	No	No	No

Chi-cuadrado

Grados de libertad	8			
Estadística	15.301			
Valor P	0.05355			
Rango	16			
?	0.2	0.1	0.05	0.02
Valor crítico	11.03	13.362	15.507	18.168
Rechazar?	Sí	Sí	No	No

Con las pruebas de bondad de ajuste (Kolmogorov-Smirnov, Anderson-Darling y Ji-Cuadrada) se corrobora que los datos siguen una distribución Weibull, pues para un alfa de 0.05 se tiene el mismo resultado en las 3 pruebas: No se rechaza la hipótesis nula, por lo tanto los datos siguen una distribución Weibull.

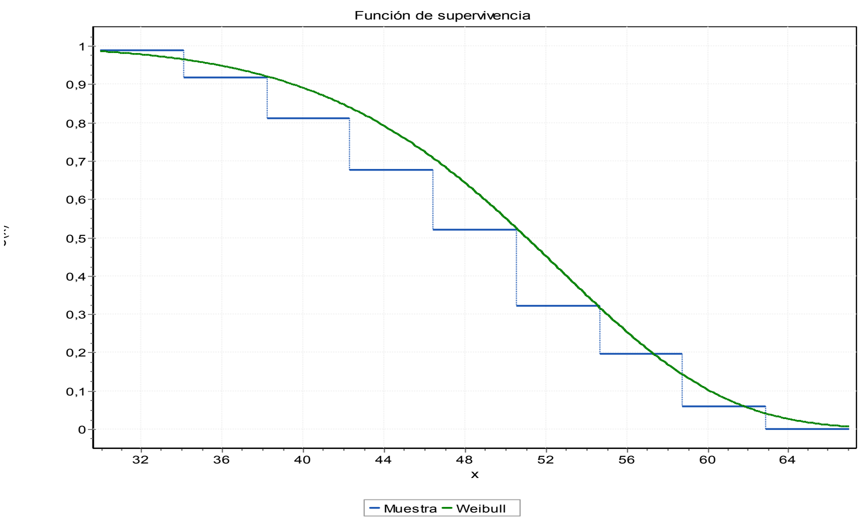
Análisis de Supervivencia para el factor Edad

Dado lo anterior se tiene que los datos siguen una distribución Weibull, por lo que en el análisis de supervivencia se hará con respecto a esta distribución: Para el análisis se siguen utilizando los asegurados que se siniestraron, o llegaron al final de la vigencia de la póliza en otras palabras los datos *No Censurados*.

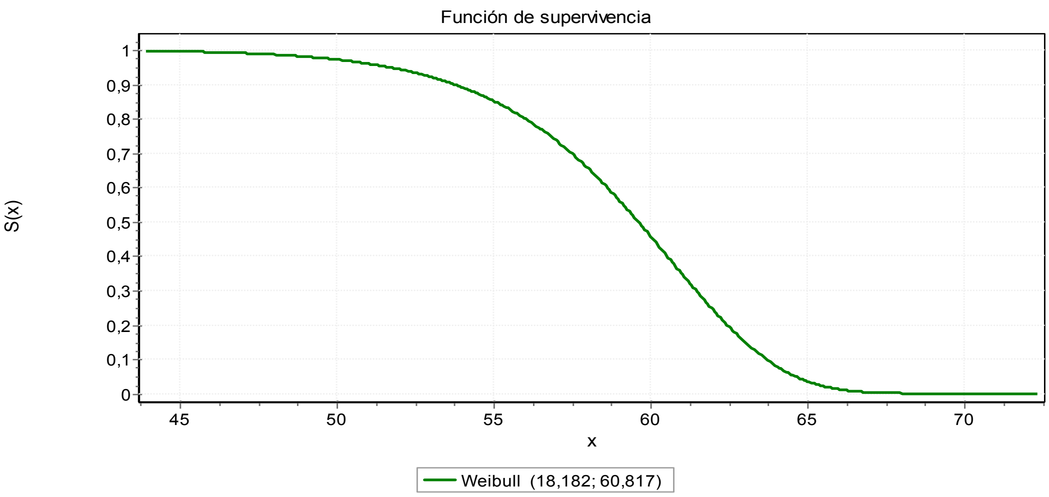
La función de supervivencia para una distribución Weibull es la siguiente:

$$S(t) = \exp[(\lambda t)^k],$$

Para los datos del análisis se tiene el parámetro λ= .018, con un índice k=7.04, que arrojan la siguiente gráfica:



Con el análisis de supervivencia, también se obtuvo que la edad promedio de supervivencia es de 51 años, este datos nos será de suma importancia pues al hacer zoom a esta población se determina que la población permanece con vida hasta la edad del fin de contrato del seguro (60 años), lo cual se puede observar en la siguiente gráfica.



Con respecto a la edad, el análisis de supervivencia determina que a la edad de 51 años, la población permanece con vida hasta la edad “n”, es decir, al final del contrato. Por lo que se crea la regla de discriminación; para que el pagare genere los rendimientos deseados se necesita *invertir en pólizas con asegurados que tengan una edad mínima de 51 años.*

Definición del instrumento financiero

Dado lo anterior, conocemos a los asegurados con los cuales se podrá hacer una inversión que genere mayores rendimientos con el instrumento financiero con una tasa de al menos 4%.

$R \quad R_2(1.04) \qquad\qquad R_{n-1}(1.04)=SA$

1 2 3 4 5 6 . . . n-1 n

dónde

R= monto de inversión inicial

R_t= monto de inversión al tiempo t que es igual al monto de inversión anterior más el rendimiento ganado en el periodo t-(t-1), i.e. R_2=R+R(1.04).

Análisis con el modelo de Markov Oculto para el Factor Fumar

```
> # Secuencia de observación para el no fallecimiento

> observation = c("0","0","0","0","0","0","0","0","0","0","1")

> # Baum-Welch

> bw = baumWelch(hmm,observation,10)

> print(bw$hmm)

$States
[1] "Fumar"    "No Fumar"

$Symbols
[1] "1" "0"

$startProbs
  Fumar No Fumar

   0.51     0.49

$transProbs
        to
from          Fumar        No Fumar

  Fumar    0.3463586   0.6536414

  No Fumar 0.1868674 0.8131326

$emissionProbs
        symbols

states            1        0

  Fumar    0.14150892 0.8584911

  No Fumar 0.07427309 0.9257269

> #Aplicando Viterbi

> viterbi = viterbi(hmm,observation)

> print(viterbi)
 [1] "No Fumar" "No Fumar" "No Fumar" "No Fumar" "No Fumar" "No Fumar"

 [7] "No Fumar" "No Fumar" "No Fumar" "No Fumar" "Fumar"

# Esta última es la secuencia más probable del comportamiento del factor de riesgo
```

V.-Síntesis y Conclusiones

En este capítulo se da a conocer la síntesis del trabajo realizado, el cual se explica en los capítulos anteriores, además de las conclusiones a las que se llegaron en el presente trabajo.

Síntesis del Trabajo

*E*l trabajo está desarrollado en base al producto de una institución aseguradora. El cual está a la venta para trabajadores, donde la edad mínima para poder adquirirlo es de 22 años y la cobertura es hasta los 60 años. También tomamos en cuenta otra característica del asegurado, tipo de hábito (Fuma o No fuma), para poder tener variables con la información completa en los asegurados y que además puedan tener cambios con respecto al tiempo, ya que el sexo de los asegurados no todos la tenían y no varía con el tiempo.

Estos datos corresponden a la base histórica de 5 años, con esta información se elaboró el análisis de supervivencia, para ello primero se realizaron las pruebas de bondad de ajuste con las que se obtuvo la función de distribución que mejor se ajustaba a los datos, con respecto a la edad de salida. Esta función de distribución fue la distribución Weibull, con ella se obtuvo la edad de supervivencia que es de 51 años. Este resultado nos ayudaría a determina que a la edad de 51 años la población permanece con vida hasta el final del contrato.

Después se realizó el análisis con base a las cadenas de Markov Ocultas. Este enfoque nos permitió modelar la dinámica de la variable “tipo de hábito” y así generar una matriz de transición que ayudaría a encontrar la probabilidad de que un asegurado fuera bueno o no. Al determinar el estatus se procede al proceso de inversión ya que dado un buen asegurado, podemos hacer una inversión que genere mayores rendimientos.

En el proceso de Inversión, nos enfocamos en las características que debería tener el pagaré para que fuese una mejor inversión. Estas son:

1.- El Capital Inicial de Inversión,

2.- El Ingreso periódico, y

3.- La tasa de inversión.

Como los clientes ya habían pasado por el primer filtro, en este paso se tiene que el Capital Inicial es bueno pues son asegurados que han podido pagar todas sus primas. Entonces nos interesa que la tasa sea mayor a la convenida en el contrato de seguro.

De esta manera se pudo determinar a los asegurados ideales y por tanto cumplir con la directiva de *"Generar mayores rendimientos a la empresa con las inversiones de los buenos asegurados".*

Conclusiones

Se concluye que el sistema para la Administración de Riesgos propuesto en el trabajo podría llegar a funcionar para un compañía de seguros, ya que nos ayuda a encontrar el grupo selecto de clientes con los cuales podemos generar una mayor rentabilidad para la empresa. Pues con los filtros que haría el sistema sabemos que los clientes llegarán a pagar en su mayoría o totalidad la prima necesaria para generar el pago de la Suma Asegurada, así entonces podremos saber que contamos con la parte correspondiente de la reserva para invertir en el instrumento financiero que nos generará mayores rendimientos, pues tenemos la certeza de que ese dinero lo podremos tener a tiempo para hacer frente a la obligación adquirida con el asegurado.

El sistema se puede adaptar a cualquier producto de una empresa, siempre y cuando se cuente con la información necesaria, en este caso, sólo se tomaron 2 factores pues no se tuvo mayor información que fuera consistente, ya que se necesitan características bien definidas para todos los asegurados y que podamos modelar según el tiempo (ver el comportamiento de la variable, por ejemplo en el hábito de fumar: No, No, No, Si).

VI.- Anexos

Códigos

Para poder correr lo que es nuestro sistema para la administración de riesgo se requiere el programa “R” (ya que es libre y tiene programadas las funciones que se requieren para el proceso de la administración de riesgos), programa estadístico que nos ayuda a generar las Redes Bayesianas, que utilizamos en el proceso, es decir, el modelo oculto de Markov (HMM), las observaciones con Baum-Welch y el algoritmo de Viterbi, que nos ayuda a encontrar las observaciones del HMM.

Código para obtener la probabilidad de un asegurado dado el tipo de hábito y la edad al querer contratar el seguro.

```
# Definicion de la Cadena de Markov Oculta para el caso de
tipo de hábito

hmm = initHMM(c("Fumar","No Fumar"),c("1","0"),

     startProbs=c(.51,.49),

     transProbs=matrix(c(.7,.3,.3,.7),2),

     emissionProbs=matrix(c(.8,.2,.2,.8),2))

print(hmm)

# Secuencia de observaciones para el no fallecimiento

observation = c("0","0","0","0","0","0","0","0","0","0","1")

# Método para ajustar los parámetros y obtener la probabilidad
para obtener la secuencia esperada y mejorar el modelo

# Baum-Welch

bw = baumWelch(hmm,observation,10)

print(bw$hmm)

#Aplicando el algoritmo de Viterbi para ver cómo debe ser el
comportamiento del cliente    para que sea un buen asegurado,
encontrar la secuencia de estados ocultos

viterbi = viterbi(hmm,observation)

print(viterbi)
```

VII.- Referencias

Bibliografía:

1 *J.M. Bernardo, *Bioestadística una perspectiva Bayesiana* , VICENS Universidad, 1a. Ed., España, 1981, pp. 1-44.

2 *José M. Bernardo y Adrian F. M. Smith, *Bayesian Theory*, John Wiley & Sons, Spain, 1994, pp. 56-67.

3 *Evgueni D. Solojentsev, *Scenario Logic and Probabilistic Management of Risk in Business and Engineering*, Springer Optimization and Its Applications, Vol.20, 2a. Ed., 1991.

4 * Díaz Carmen, *Futuros y opciones sobre futuros financieros: teoría y práctica*, México, Prentice Hall, 1998, pp.125-171.

5 * Diaz Tinoco Jaime, Fausto Hernandez Trillo, *Futuros y opciones financieras: una introducción* , 3a. Ed. México, Limusa, Noriega, 2000.

6 *Elisa T. Lee, John Wenyu Wang, *Statistical Methods for Survival Data Analysis*, Wiley series in probability and statistics, 3a.Ed., 2003.

7 *Casuality Actuarial Society*, Foundations of casuality Actuarial Science*, United Book Press Inc. 4a ed, USA 2001.

8 *Nir Friedman, Kevin Murphy & Stuart Rusell, *Learning the Structure of Dynamic Probilistic Networks,* Uncertainty Artificial Intelligence, 1999, pp.139-147.

9 *Stephenson Todd A., *An Introduction To Bayesian Network Theory and Usage,* IDIAP Research Report 00-03, February 2000.

10 *Gahramani Z. *An Introduction to Hidden Markov Models and Bayesian Networks,* 2001.

11 *Dempster A., Laird N., Rubin D., *Maximum Likelihood from Incomplete Data via the EM Algorithm.,*1977.

12 * Nota técnica del producto dotal de la empresa.

Páginas de Internet consultadas:

13 *http://www.ehowenespanol.com/significado-suscripcion-seguro-hechos_361283/

14 *http://www.theirm.org/publications/documents/rm_standard_spanish_15_11_04.pdf

15 *http://www.economia.com.mx/origen_y_mision_del_seguro.htm

16 *http://www.gestiopolis.com/canales2/finanzas/1/admonriego.htm

17 *http://www.matematicas.unam.mx/actuaria/plan2006/Administracion%20Actuarial.pdf

18*http://www.cnsf.gob.mx/AcercadelaCNSF/Documents/GES-04_CONSTITUCION%20DE%20RESERVAS%20APROBACIN.pdf

19 *http://www.utm.mx/edi_anteriores/temas45/1ENSAYO_45_2.pdf

20 * http://www.mathwave.com/help/easyfit/html/analyses/fitting/manual.html

21*http://raudys.com/kursas/Options,%20Futures%20and%20Other%20Derivatives%207th%20John%20Hull.pdf Capítulo 11