

**¿Qué prueba o secuencia de pruebas
es óptima para determinar una
condición de salud?**

Yanink Neried Caro Vega

Director de Tesis
Dr. José Andrés Christen Gracia

8 de agosto de 2005

Agradecimientos

El escribir mi tesis de maestría, tras casi tres años de llegar a este país, es una meta que varias ocasiones pensé sería imposible de alcanzar. Me siento profundamente agradecida con todas las personas que contribuyeron al logro de este sueño.

Gracias a mi familia, que ha sabido acompañarme y apoyarme a pesar de la distancia y las dificultades que implican mi ausencia. Su cariño y confianza me han dado la fuerza y la seguridad que tantas veces perdí.

A los amigos que encuentre, que me hicieron sentir como en casa, me reconfortaron y animaron a continuar. Especialmente a quienes me acogieron en su vida y me hicieron parte de ella. Les debo muchos momentos felices y valiosas lecciones que van más allá de la estadística y la matemática. A mis viejos amigos, que siempre han estado conmigo aún cuando me equivoco o me alejo.

A mis maestros, sus enseñanzas, paciencia y exigencia permitieron que llegara a este momento. Quiero mencionar especialmente a la Dra. Graciela González Farías a quien nunca terminaré de agradecerle el darme la oportunidad de venir a México y al Dr. Andrés Christen, mi “profe”, quien me guió en el proceso de elaboración de la tesis y me dió la confianza y herramientas necesarias para trabajar en un tema aplicado como siempre quise. Para terminar debo agradecer a CIMAT, su apoyo económico y la colaboración de sus funcionarios fueron parte importante de mi sobrevivencia en la Maestría.

A todos de nuevo muchas gracias, sin ustedes jamás lo habría conseguido.

Yanin

Índice general

1. Introducción	1
1.1. Notación	6
2. Conceptos Básicos	7
2.1. Teoría de utilidad	7
2.2. Árboles de decisión y Backward Induction	9
2.2.1. Demostración	13
2.3. Conjuntos de funciones de utilidad	15
3. Utilidades en Medicina	21
3.1. Revisión Bibliográfica	21
3.1.1. QALYs: Quality Adjusted Life Years	22
3.1.2. Otras alternativas	24
3.2. ¿Por qué utilizar QALYs?	25
3.3. Cálculo de utilidades usando QALYs	27
3.4. QALYs vs. costos	28
4. Solución del problema	33
4.1. Cálculo de probabilidades	33
4.2. Solución al problema con una sola función de utilidad	36
4.2.1. Ejemplo de la aplicación del algoritmo de <i>Backward Induction</i>	39
4.3. Ejemplos y Discusión	44
4.3.1. Ejemplo 1. Estrategia III UNAIDS/WHO	44
4.3.2. Ejemplo 2. Pruebas de Monitoreo de SIDA	47
4.4. Conjuntos de funciones de utilidad	51
4.5. Recomendaciones	54

5. Conclusiones	57
A. Anexos	61
A.1. Ejemplo de aplicación del algoritmo en la Sección 4.2.1.	61
A.2. Salida completa del Ejemplo de aplicación de la Sección 4.2.1.	62
B. Anexos	63
B.1. Uso del programa	63

Capítulo 1

Introducción

Desde 1985, las pruebas del VIH han sido esenciales para garantizar la seguridad de los suministros de sangre, para vigilar el avance de la epidemia y para diagnosticar los casos de infección por el virus.

ONUSIDA, 1997

La práctica común de los médicos de diagnosticar teniendo en cuenta su experiencia previa, el historial clínico y los antecedentes familiares de un paciente, es natural y ampliamente aceptada. Este comportamiento natural corresponde al enfoque bayesiano de considerar probabilidades y distribuciones *a priori* para estimar parámetros desconocidos.

Adicionalmente, en el campo de la salud, la teoría de decisiones es una rama que cobra cada vez mayor importancia. Los recursos para la salud son bajos en la mayoría de los países y su administración debe optimizarse. Por tal razón, la aplicación de nuevos tratamientos y adquisición de material de laboratorio y nuevas tecnologías se debe respaldar con estudios de costo-efectividad que garanticen la obtención de máximos beneficios.

En este contexto, las utilidades son medidas numéricas que representan el valor que tiene para un individuo un estado de salud y lo deseable que le resulta. Como las preferencias de un individuo condicionan su decisión de someterse a un tratamiento que le puede implicar mejoría y riesgo simultáneamente, se puede analizar este problema dentro de la perspectiva bayesiana de maximización de utilidades en teoría de decisión.

El interés de este trabajo es proponer una metodología apoyada en la teoría de decisión que permita definir la prueba o secuencia de pruebas óptimas a emplear para determinar la presencia de una condición de salud en una persona.

Para ilustrar el problema y presentar una aplicación práctica de su solución, se presentan ejemplos sobre el uso de las pruebas de diagnóstico de VIH (Virus de Inmunodeficiencia Humana) y monitoreo del SIDA (Síndrome de Inmunodeficiencia Adquirida) para determinar la presencia del virus en una persona o hacer seguimiento de su enfermedad.

En el primer contexto, para diagnosticar el VIH, que causa el SIDA, se utilizan varias pruebas que dependiendo de la prevalencia de la enfermedad en una región determinada e inclusive del momento de su aplicación, difieren en sus resultados. Debido a esto un positivo inicial debe ser confirmado por la misma prueba o una de mayor especificidad. UNAIDS/WHO (1997) reconoce la importancia de las pruebas diagnósticas para frenar la epidemia, prevenir oportunamente su expansión y reducir costos de tratamiento. Para ello propone una serie de estrategias de aplicación de pruebas diagnóstico considerando el objetivo de la prueba pero sin aparente justificación estadística. (Ver Figura 4.4 en la Sección 4.3.1).

La propuesta a resolver en este contexto considera tres tipos de pruebas:

- Pruebas ELISA
- Pruebas Rápidas
- Western Blot

Las pruebas difieren en su costo y condiciones operativas de sensibilidad y especificidad, además de aspectos técnicos como su facilidad de aplicación, requerimiento de equipo especializado y cantidad de muestras que pueden ser procesadas a la vez DELIVER (2002). Por supuesto, también las marcas comerciales de las pruebas intervienen en estas diferencias, pero estos hechos no se tendrán en cuenta en este trabajo. El modelo presentado se elabora con condiciones básicas de información, la actualmente disponible, que simplifica el problema y su aplicación. Sin embargo el modelo cuenta con la flexibilidad suficiente para incluir más detalles en desarrollos posteriores.

Para el segundo problema, se sabe que el virus VIH ataca el sistema inmunológico, específicamente a las células de defensa CD4, por lo que una variación de la cantidad normal de estas células define el inicio del SIDA y el nivel de enfermedad en una persona. Como la forma de monitorizar el avance de la enfermedad o la efectividad de un tratamiento tiene diversos niveles o pruebas, en este trabajo se presentará el caso de cuatro de estas posibles pruebas, ordenadas en forma ascendente de acuerdo a la confiabilidad de sus resultados y a su costo. Se consideran entonces:

- Índice de Masa Corporal
- Linfocitos totales
- CD4
- Carga Viral

Generalmente se realizan distintos tipos de pruebas, combinaciones o secuencias de ellas; para determinar el estado de un paciente y definir el momento en que debe iniciarse un tratamiento o modificarlo. Pero no existe una estrategia definida de acción, como la propuesta por UNAIDS/WHO(1997) para las pruebas diagnósticas.

El momento en que se decide iniciar o cambiar un tratamiento, es clave en el progreso de los pacientes. En países con presupuestos de salud altos, se fijan puntos de corte altos. Esto permite iniciar el tratamiento más pronto, lo que mejora la situación inmunológica del paciente. Considerar además las condiciones tecnológicas y económicas, almacenamiento, administración y análisis, hacen que el costo de las pruebas tenga implicaciones importantes en los presupuestos de salud.

Se acostumbra a iniciar un tratamiento antirretroviral a partir de un conteo de CD4 “inferior a 250”. La prueba de CD4 es bastante costosa por lo que el umbral de inicio de tratamiento varía de acuerdo al presupuesto disponible. De igual manera se usa la carga viral para determinar el inicio o cambio de tratamiento; o se combinan las dos informaciones, CD4 y carga viral, con el mismo objeto. La propuesta del Índice de Masa Corporal resultaría benéfica en sentido económico, ya que se trata sólo de la relación entre peso y altura de una persona (*peso en Kg/(altura en m)²*), sin requerir

equipo especial, mantenimiento o personal experimentado como las otras.

Bajo este contexto, el objetivo es presentar un modelo que sirva para decidir qué prueba o secuencia de pruebas usar en un número fijo de aplicaciones para diagnosticar el estado inicial de un paciente o su respuesta a determinado tratamiento.

Con esta finalidad, en los dos contextos, se requiere conocer el costo de cada una de las pruebas mencionadas así como sus probabilidades de falsos positivos y falsos negativos. Además se debe fijar una probabilidad a priori de poseer la enfermedad, conocida como prevalencia de la enfermedad, para las pruebas diagnóstico y la probabilidad a priori de responder positivamente a un tratamiento, para las pruebas de monitoreo. También datos de QALYs *Quality Adjusted Life Years* (ver Capítulo 3), que básicamente son utilidades en términos de salud asociadas al uso de los tratamientos posibles.

Aunque existe interés por contar con estrategias de acción costo-efectivas sobre la aplicación de pruebas de diagnóstico y monitoreo de la enfermedad, los estudios realizados hasta el momento aún no son concluyentes como se muestra en la Sección 4.3.2 y en la práctica la aplicación de pruebas se ve limitada a los recursos disponibles y a las recomendaciones generales establecidas por organizaciones como ONUSIDA/WHO(1997).

La propuesta presentada es una solución teórica al problema. El definir estrategias de aplicación de pruebas, depende de la disponibilidad de datos reales y de las circunstancias particulares de las entidades encargadas de generar las políticas de atención de pacientes con VIH/SIDA. Para ilustrar la aplicación de la propuesta se incluyen ejemplos, utilizando valores arbitrarios para los QALYs, cuya solución se obtuvo de un programa computacional que implementa el cálculo de la secuencia óptima de pruebas a aplicar para conocer el estado de un paciente.

Para resolver el problema y garantizar la comprensión de la solución propuesta, se presentan en el Capítulo 2 conceptos básicos relacionados con la teoría de la utilidad desde un punto de vista estadístico bayesiano, los elementos de un problema de decisión y su representación a través de un árbol de decisión y el procedimiento *Backward Induction*. Este procedimiento permite elegir una serie de acciones que garantizan la mayor utilidad esperada

posible. Una demostración informal de este hecho se incluye en dicho Capítulo para justificar su elección en la solución del problema central del trabajo. Finalmente se hace una presentación de cómo debería manejarse el problema para conjuntos de funciones de utilidad.

En el Capítulo 3, se hace una revisión bibliográfica del uso de las utilidades en medicina, los métodos de cálculo de utilidades, el concepto de QALY y la justificación de su elección en este contexto, además de su uso y manejo al relacionarlo con costos, introduciendo el concepto de cocientes de costo-efectividad.

El Capítulo 4 presenta la solución a los dos problemas, fijando un número límite de aplicaciones de pruebas, usando una sola función de utilidad y luego utilizando conjuntos de funciones de utilidad. Se presenta la deducción de las probabilidades implicadas en la solución del árbol, un ejemplo detallado de la solución al problema empleando Backward Induction y que incluye el cálculo de cocientes de costo-efectividad para definir el tratamiento que garantiza mayores utilidades por unidad de dinero invertido. Se presentan además aplicaciones concretas del algoritmo para el caso de una sola función de utilidad y se discuten las posibilidades e inconvenientes de emplear el algoritmo cuando se tienen conjuntos de funciones de utilidad.

Los ejemplos que aparecen en el Capítulo 4 como solución a los problemas, emplean una parte de la información de entrada real. Esto respecto a las condiciones operativas de las pruebas y costos de pruebas y tratamientos. Pero la parte referente a los QALYs no proviene de datos reales basados en estudios específicos; representa únicamente valores posibles de utilidad medidos en este esquema. Para su solución se implementó un programa en *C++* que resuelve el árbol de decisión mediante el algoritmo de Backward Induction y proporciona una secuencia óptima de pruebas a aplicar y tratamiento a seguir bajo condiciones dadas.

Finalmente el Capítulo 5 reúne las conclusiones del trabajo, menciona los avances logrados, su importancia y limitaciones y discute las posibilidades de mejora para posibles estudios posteriores. En los anexos, por otra parte, se incluye la salida del ejemplo de aplicación del algoritmo y una explicación del uso del programa.

1.1. Notación

En términos generales la notación a emplear es:

- $P(X = 1)$ Probabilidad a priori de poseer una condición E de interés.
- c_t costo de aplicar la prueba t con $t = 1, 2, \dots, T$.
- q_{jk} QALY (utilidad) de aplicar el tratamiento $j = 0, 1$ cuando la situación real de la persona con relación a la condición E , es $k = 1$ cuando se tiene la condición y $k = 0$ en caso contrario.
- $P(Y^t = 1|X = 0)$ probabilidad de falsos positivos de la prueba t
- $P(Y^t = 0|X = 1)$ probabilidad de falsos negativos de la prueba t .

En el ejemplo de las pruebas diagnóstico: la condición de salud E es la posible presencia del virus en la persona, con $k = 1$ cuando tiene el virus y $k = 0$ si no lo tiene. $P(X = 1)$ representa la probabilidad *a priori* de que una persona esté enferma, equivalente a la prevalencia de la enfermedad en la población o antecedentes de la persona. $T = 3$ porque se usan tres pruebas *Elisa*, *Rápidas* y *Western Blot* y $j = 0, 1$ significa la posibilidad de aplicar dos tratamientos dependiendo del resultado: Se asume que la persona es seropositiva, portadora del VIH si $j = 1$ o no lo es, $j = 0$.

Capítulo 2

Conceptos Básicos

2.1. Teoría de utilidad

Formalmente, un problema de decisión contempla tres elementos, un conjunto de acciones disponibles A , un conjunto de eventos inciertos E y un conjunto de consecuencias Z , relacionadas con los posibles eventos de cada acción (Bernardo y Smith, 1994 p.16). Asociado a E , existe un conjunto de parámetros T que caracterizan la realidad no conocida y modelan su incertidumbre. De allí, una acción se describe por el conjunto de distribuciones de probabilidad sobre las salidas que ésta da bajo cada posible valor del parámetro (Parmigiani, 2002 p.53).

El conjunto de consecuencias presenta un orden natural determinado por las preferencias de los individuos. En Z se pueden definir consecuencias extremas, z_* y z^* , la peor y mejor consecuencia respectivamente. Luego toda consecuencia z en Z será preferible a z_* pero menos preferible que z^* .

Se usa la notación: $p(z|a, \theta)$ para describir la probabilidad de obtener la consecuencia z al aplicar la acción a bajo un valor específico del parámetro θ . Además, como una consecuencia $z = z_{ij}$ depende de un evento E_j y de una acción a_i , se acostumbra a usar $u(z_{ij}) = u(a_i, E_j)$ (Christen, 2004 p.49) donde $u(z_{ij})$ representa la utilidad de la consecuencia z_{ij} si se toma la acción a_i y sucede el evento E_j , y $p(\theta)$ corresponde a la probabilidad a priori del parámetro θ .

La utilidad esperada de una acción a , considerando la utilidad obtenida en todas las consecuencias y los valores posibles del parámetro θ , se define entonces como:

$$U^*(a) = \int_Z \int_{\theta} u(z)p(z|a, \theta)p(\theta)d\theta dz. \quad (2.1)$$

En algunos casos es útil manejar una función de pérdida $L(z)$ en vez de una función de utilidad, luego se tiene la relación, $L(z) = -u(z)$.

Interesa ahora elegir la acción que reporte mayor utilidad esperada. Se busca la acción óptima a^* que resulta al maximizar $U^*(a)$, es decir a^* en A tal que $U^*(a^*) = \max_{a \in A} U^*(a)$. El siguiente ejemplo será de utilidad para aclarar los conceptos referidos.

Suponga que un paciente con VIH tiene la opción de evaluar su condición para que se le recomiende un tratamiento específico y que tiene una probabilidad de 0.7 de tener la condición evaluada. Si tras la prueba se le recomienda seguir el tratamiento A tendrá una ganancia en términos de años de buena salud de 10, en caso de que lo requiera realmente, es decir, si el paciente realmente alcanzó una condición grave de su enfermedad. Si no es así, la prueba arrojó un falso positivo, y se obtendrá un beneficio de 0. Si por el contrario se recomienda el tratamiento B, y la persona realmente se encuentra en un nivel avanzado de la enfermedad obtendrá una utilidad de 0 y en otro caso de 7 años de buena salud. Lo anterior implica que el tratamiento A es mejor cuando la persona realmente requiere tratamiento y el B es mejor cuando la persona aún no alcanza el nivel crítico. La Tabla 2.1 representa el esquema de utilidades.

	$X = 0$	$X = 1$
Tto A	0	10
Tto B	7	0

Tabla 2.1: Ejemplo de tabla de utilidades.

Se puede considerar el problema descrito como un problema de decisión. El conjunto de acciones está dado por: $A = a_1, a_2$; a_1 es aplicar el tratamiento A y a_2 proporcionar el tratamiento B. Los eventos inciertos son dos, E_1 la persona alcanzó el nivel superior de la enfermedad y E_2 su condición es

estable. El conjunto de consecuencias se refiere a las respuestas en salud posibles determinadas por las acciones y los eventos correspondientes. Así, z_1 representa el resultado cuando la persona toma el tratamiento A y estaba enferma realmente, z_2 cuando la persona se sometió al tratamiento A pero no lo requería. Y z_3 si la persona siguió el tratamiento B y realmente estaba enferma ó z_4 cuando tomó el tratamiento B pero no estaba enferma. Asociadas a estas consecuencias, se consideran las utilidades $u(z_1) = u(A, X = 1) = q_{00} = 0$, $u(z_2) = u(A, X = 0) = q_{01} = 10$, $u(z_3) = u(B, X = 1) = q_{10} = 7$ y $u(z_4) = u(B, X = 0) = q_{11} = 0$.

Evaluar la utilidad esperada de la acción a_1 corresponde a:

$$U^*(a_1) = 0.7 * 10 + 0.3 * 0 = 7,$$

y respectivamente para la acción a_2 esto es

$$U^*(a_2) = 0.7 * 0 + 0.3 * 7 = 2.1.$$

En este caso la acción a_1 maximiza la utilidad esperada. Es evidente que la mejor opción es prescribir el tratamiento A.

2.2. Árboles de decisión y Backward Induction

Un árbol de decisión es una representación gráfica de un problema de decisión. La estructura tiene los elementos básicos definidos anteriormente. Un conjunto de acciones A , un conjunto de eventos inciertos E y un conjunto de consecuencias Z . La Figura 2.1 representa gráficamente las acciones posibles a elegir, los eventos que pueden ocurrir y las consecuencias obtenidas en cada caso. Los cuadrados simbolizan *nodos de decisión*, implican que una acción se debe elegir. Los círculos son *nodos aleatorios* representan la parte aleatoria del problema, son los posibles eventos que guían a través de las ramas del árbol a las consecuencias consideradas.

Para simplificar el problema, y explicar el modelo empleado en la implementación del programa, se tomarán las siguientes restricciones:

- Sólo se aplicarán las pruebas 1 y 2.

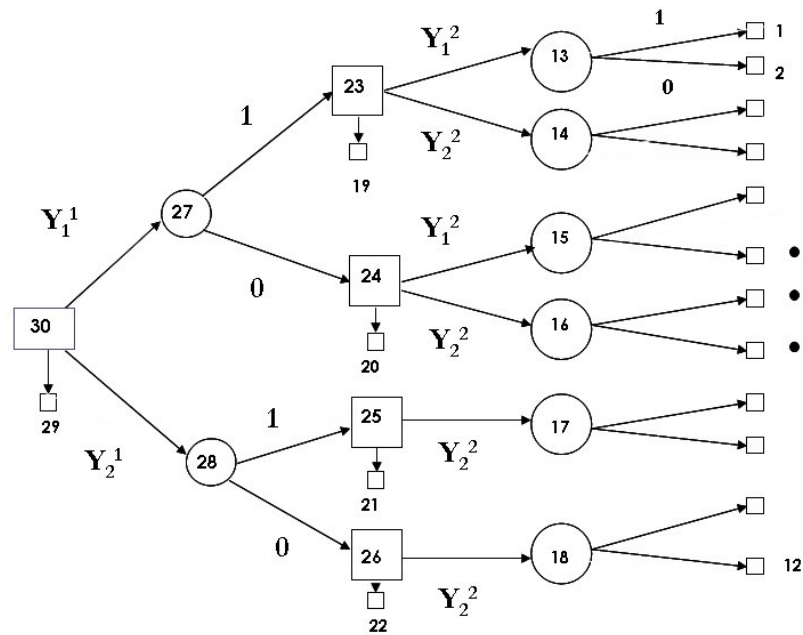


Figura 2.1: Árbol de decisión

- Sólo se aplicará dos pruebas en total a un paciente.
- Sólo es permitido escalar en las pruebas. Pasar de la prueba 1 a la 2 o mantenerse en la prueba anterior.

Además se usará la siguiente notación:

- $P(Y_1^1 = 1)$ Probabilidad de que la prueba 1 resulte positiva en la aplicación 1. Existe un umbral a partir del cuál se determina que la persona está enferma o que no ha mejorado en su tratamiento. En forma equivalente $P(Y_1^2 = 1)$ para la prueba 2 y $P(Y_1^2 = 0)$ cuando la prueba resulta negativa.
- $P(X = 1|Y_1^1 = 1)$ Probabilidad de que la persona realmente tenga la condición E dado que la prueba 1 en la aplicación 1 resulta positiva.
- $P(Y_2^1 = 1|Y_1^1 = 1)$ Probabilidad de que la prueba 1 en la aplicación 2 resulte positiva dado que la prueba 1 en la aplicación 1 fue positiva.

- $P(X = 1|Y_2^1 = 1, Y_1^1 = 1)$ Probabilidad de que la persona realmente tenga la condición E dado que las dos aplicaciones de la prueba 1 lo afirman.
- $\theta = P(X = 1)$.
- q_{01} , q_{00} son las utilidades del tratamiento 0 para el caso en que la persona realmente se encuentra en la condición 1 ó en la condición 0 respectivamente. Equivalentemente se usa q_{11} y q_{10} para el tratamiento 1.

El árbol de la Figura 2.1 representa el problema simplificado. En él se advierte cada camino como una secuencia de pruebas a aplicar y las dos salidas posibles de la prueba. Por ejemplo, la secuencia de la primera rama superior del árbol sugiere: la primera aplicación de la primera prueba, Y_1^1 en seguida la posibilidad de obtener un 1, aplicar por segunda vez la prueba 1 y obtener el resultado 1.

El primer nodo del árbol *nodo30* es de decisión; se refiere a la posibilidad de decidir cuál prueba aplicar o recomendar uno de los tratamientos, sin hacer pruebas. Tras la aplicación de cada prueba y la obtención de la respuesta, nuevamente aparecen nodos de decisión, para responder a la pregunta: ¿Se aplicará la prueba 1 ó 2, o un tratamiento? Los nodos 29, 19, 20, 21, 22 corresponden al caso en que se decide no realizar más pruebas sino considerar la información obtenida hasta al momento y aplicar un tratamiento.

En la parte superior del árbol, determinada por la rama correspondiente a la aplicación de la primera prueba Y_1^1 y las ramas que se derivan de ella, se tienen tres acciones: aplicar la prueba 1, aplicar la prueba 2 y detener el experimento recomendando un tratamiento. En la parte inferior, cuando se aplica primero la segunda prueba Y_1^2 , se tienen sólo las opciones a_2 aplicar la prueba 2 y a_3 parar la secuencia de pruebas y recomendar un tratamiento. Los eventos aleatorios son E_1 la prueba resulta positiva y E_2 la prueba es negativa. Las utilidades de los tratamientos son q_{01} , q_{00} , q_{11} , q_{10} y las utilidades de las secuencias de acciones son *nodo1*, ..., *nodo12*.

El procedimiento de *Backward Induction* permite elegir la mejor secuencia de acciones a tomar a lo largo de un árbol de decisión finito (Parmigianni, 2002 p.173 y Bernardo y Smith, 1994 p.59). Esto se logra al calcular las

utilidades en cada conjunto posible de acciones y escoger la de utilidad máxima. El procedimiento se realiza en sentido inverso al de construcción del árbol es decir de derecha a izquierda. Al llegar a la raíz del árbol se tiene completamente determinando el mejor camino posible para recorrerlo y obtener las utilidades máximas.

Para aplicar el algoritmo de Backward Induction al árbol de la Figura 2.1, se deben considerar las utilidades asociadas a las consecuencias z_1, \dots, z_{12} a todas las secuencias posibles, esto es *nodo1*, ..., *nodo12*. Se calcula la utilidad de a_1 como el máximo de los valores esperados de esta acción, considerando las utilidades ofrecidas por los dos tratamientos. Es decir:

$$\begin{aligned} \text{nodo1} = \max\{ & P(X = 1|Y_2^1 = 1, Y_1^1 = 1)q_{01} + P(X = 0|Y_2^1 = 1, Y_1^1 = 1)q_{00}, \\ & P(X = 1|Y_2^1 = 1, Y_1^1 = 1)q_{11} + P(X = 0|Y_2^1 = 1, Y_1^1 = 1)q_{10}\}. \end{aligned}$$

Este es, el cálculo del nodo 1 que corresponde al caso en que se aplicó en dos ocasiones la prueba 1 y resultó positiva. El nodo 2 en forma análoga corresponde al caso en que la segunda aplicación de la primera prueba resulta negativa y la primera fue positiva:

$$\begin{aligned} \text{nodo2} = \max\{ & P(X = 1|Y_2^1 = 0, Y_1^1 = 1)q_{01} + P(X = 0|Y_2^1 = 0, Y_1^1 = 1)q_{00}, \\ & P(X = 1|Y_2^1 = 0, Y_1^1 = 1)q_{11} + P(X = 0|Y_2^1 = 0, Y_1^1 = 1)q_{10}\}. \end{aligned}$$

nodo13 se calcula tomando el valor esperado de las posibles salidas al aplicar la prueba 1 por segunda vez, dado que la prueba 1 en el primer intento resultó positiva, considerando las utilidades *nodo1* y *nodo2*, esto es:

$$\text{nodo13} = P(Y_2^1 = 1|Y_1^1 = 1)\text{nodo1} + P(Y_2^1 = 0|Y_1^1 = 1)\text{nodo2}.$$

nodo14 corresponde a la utilidad esperada de aplicar la prueba 2 después de la prueba 1, esto es:

$$\text{nodo14} = P(Y_2^2 = 1|Y_1^1 = 1)\text{nodo3} + P(Y_2^2 = 0|Y_1^1 = 1)\text{nodo4}.$$

nodo19 es el máximo entre $P(X = 1|Y_1^1 = 1)q_{01} + P(X = 0|Y_1^1 = 1)q_{00}$ y $P(X = 1|Y_1^1 = 1)q_{11} + P(X = 0|Y_1^1 = 1)q_{10}$ esto corresponde a tomar la decisión sin realizar más pruebas.

$nodo23$ es el máximo de las utilidades calculadas en los extremos de sus ramas, $\max\{nodo19, nodo13, nodo14\}$ e indica el camino a seguir, por ejemplo si $nodo14$ resulta mayor se recomienda aplicar la prueba 2 en la segunda aplicación dado que la primera resultó positiva.

En forma analóga $nodo24$ corresponde al máximo entre las utilidades esperadas de: seguir la secuencia de pruebas cuando la primera del tipo 1 resultó negativa y detener la aplicación de pruebas.

$nodo27$ es la utilidad esperada de aplicar la prueba 1 en el primer intento. $nodo27 = P(Y_1^1 = 1)nodo23 + P(Y_1^1 = 0)nodo24$ y $nodo28 = P(Y_1^2 = 1)nodo25 + P(Y_1^2 = 0)nodo26$ es la utilidad esperada de aplicar la prueba 2 en el primer intento.

Finalmente, $nodo29$, que corresponde al caso en que se decide no realizar ninguna prueba, es el máximo entre las utilidades obtenidas a través de los dos tratamientos y con la probabilidad a priori de tener la condición de interés. Esto es el máximo entre $P(X = 1)q_{01} + P(X = 0)q_{00}$ y $P(X = 1)q_{11} + P(X = 0)q_{10}$. De ahí se calcula $nodo30$ que corresponde a la mejor decisión a tomar respecto a la secuencia de pruebas a seguir: $nodo30 = \text{Max}\{nodo27, nodo28, nodo29\}$.

2.2.1. Demostración

A continuación se hará un esbozo de demostración, basado en la efectuada por Bernardo y Smith (1994, p.59-60), para verificar que el procedimiento de Backward Induction encuentra la utilidad máxima. Inicialmente se hará sin considerar la acción de parar, esto es para un problema general de decisión entre acciones secuenciales.

Sea $A = \{a_i, i \in I\}$ el conjunto de las acciones posibles en el árbol de decisión de la Figura 2.2 para cada a_i existe una clase de eventos aleatorios $\{E_{ij}, j \in J_i\}$ y $\{c_{ij}, j \in J_i\}$ un conjunto de consecuencias asociado a cada evento resultante de la acción a_i . Sea n el número de etapas de decisión del árbol determinado por los nodos de decisión. El procedimiento sugiere definir la decisión de la primera etapa, para ello se debe encontrar la acción que maximice la utilidad en la segunda etapa y así sucesivamente hasta llegar a la última etapa.

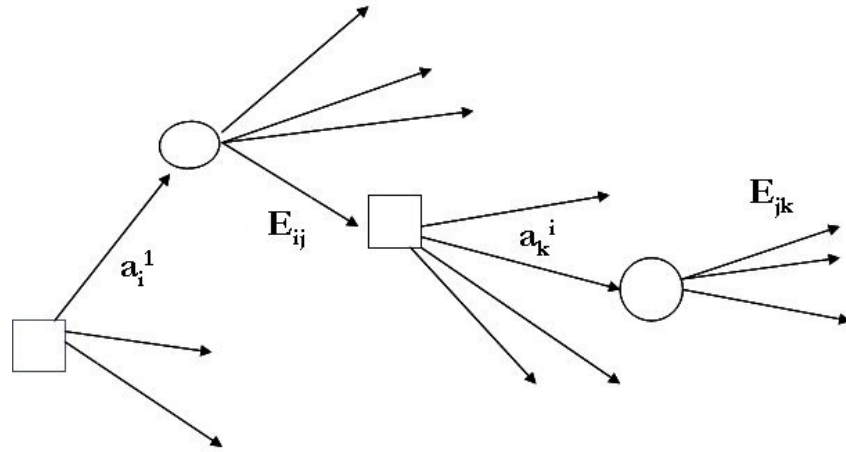


Figura 2.2: Árbol de decisión general

Las acciones de la primera etapa se pueden escribir como:

$$a_i^{(1)} = \{max_{k \in K_{ij}} a_k^{(2)} | E_{ij}, j \in J_i^{(1)}\},$$

es decir, considerando la mejor acción de la segunda etapa y dada la información correspondiente a la acción i en la primera etapa. En forma análoga las acciones de la etapa 2 dependen de la máxima acción de la etapa 3 y así sucesivamente. En general el conjunto de acciones de la etapa m para $m = 1, 2, \dots, n - 1$ se puede escribir como:

$$a_i^{(m)} = \{max_{k \in K_{ij}} a_k^{(m+1)} | E_{ij}, j \in J_i^{(m)}\},$$

y

$$a_i^{(n)} = \{c_{ij} | E_{ij}, j \in J_i^{(n)}\}$$

es el conjunto de acciones de la última etapa basada en las consecuencias de estas acciones.

Se sabe que una acción a es preferida a otra acción a' si la utilidad esperada de la primera es mayor que la utilidad esperada de la segunda. De esta manera elegir la acción óptima en la etapa m equivale a calcular la

utilidad esperada en esa etapa:

$$U^*(a_i^{(m)}) = \sum_{j \in J_i^{(m)}} \max_{k \in K_{ij}} U^*(a_k^{(m+1)}) P(E_{ij}).$$

$$U^*(a_i^{(n)}) = \sum_{j \in J_i^{(n)}} U(c_{ij}) P(E_{ij}).$$

Es decir que si se sabe el valor de las utilidades máximas en la etapa $m+1$, se pueden calcular las utilidades máximas en la etapa m . De esta manera se pueden calcular las utilidades en el horizonte e “inducir regresivamente” para calcular las utilidades y la decisión óptima al inicio del árbol que es lo que se busca.

Luego, el cálculo de la utilidad esperada en la etapa m depende de la máxima utilidad esperada de la etapa $m+1$ y así hasta la etapa $n-1$ que depende de maximizar las utilidades de las consecuencias de la etapa n . Aplicar el algoritmo requiere entonces maximizar la utilidad esperada en la etapa n , luego resolver la etapa $n-1$ para maximizar la utilidad esperada condicional, es decir para hacer la elección óptima basada en la n -ésima etapa y continuar así hasta alcanzar la primera etapa.

Al incluir la acción de parar en cada nodo de decisión, se debe calcular su utilidad maximizando la posibilidad de aplicar un tratamiento dada la información disponible hasta el momento. (Ver Sección 2.2.) Luego basta comparar la utilidad de parar con la máxima utilidad esperada de seguir con la acción óptima y decidir a favor de la utilidad más alta.

2.3. Conjuntos de funciones de utilidad

Tradicionalmente se cuenta con una sola función de utilidad dada por el valor de utilidad correspondiente a una consecuencia específica. Se llama función de utilidad a una aplicación del conjunto de consecuencias al conjunto de números reales, (Ver Bernardo y Smith, 1994 p.50)

$$u : Z \rightarrow R.$$

En ocasiones no es posible definir una función única y se cuenta con conjuntos de funciones de utilidad. Aunque el problema de decisión

sigue siendo definir cuál tratamiento recomendar, es necesario implementar un procedimiento que permita evaluar cada vez, considerando todas las utilidades posibles del conjunto propuesto, el resultado de una acción tomada y decidir continuar su aplicación o desecharla.

Para este caso, se propone considerar las utilidades q_{ij} como valores indexados por $\alpha \in [0, 1]$ y determinados por los valores de QALY mínimo $q_m(T_i, E_j)$ y máximo $q_M(T_i, E_j)$ de cada tratamiento y condición real. Estas son las utilidades mínima y máxima esperada de cada tratamiento. Así:

$$\begin{aligned} q_{01}^\alpha &= \alpha q_M(T_0, E_1) + (1 - \alpha)q_m(T_0, E_1), \\ q_{00}^\alpha &= \alpha q_M(T_0, E_0) + (1 - \alpha)q_m(T_0, E_0), \\ q_{11}^\alpha &= \alpha q_M(T_1, E_1) + (1 - \alpha)q_m(T_1, E_1), \quad y \\ q_{10}^\alpha &= \alpha q_M(T_1, E_0) + (1 - \alpha)q_m(T_1, E_0). \end{aligned}$$

Para ejemplificar este caso se considerará una variación del ejemplo descrito en la Sección 2.1. En el mismo contexto, el médico cuenta con rangos de utilidades, en términos de QALYs. La Tabla 2.2 muestra los conjuntos de QALYs para esta situación.

	$X = 0$	$X = 1$
Tto 1	[4,9]	[2,8]
Tto 2	[6,7]	[5,7]

Tabla 2.2: Conjuntos de QALYs

De igual manera interesa definir la acción óptima a seguir por lo que interesa calcular los QALYs esperados de cada tratamiento de acuerdo a la información disponible. Tomar la decisión en este caso requiere un poco más de trabajo. Se deben considerar elementos como los QALYs mínimos esperados, los máximos y el rango que define cada acción.

La Figura 2.3(a), de valores de α vs. valores de QALYs, muestra el comportamiento de los conjuntos de utilidades. En este ejemplo se considera $P(X = 1) = 0.7$. Se observa que el tratamiento 2 es superior al tratamiento 1 cuando $X = 1$ para valores inferiores de $\alpha = 0.75$ y cuando $X = 0$ para

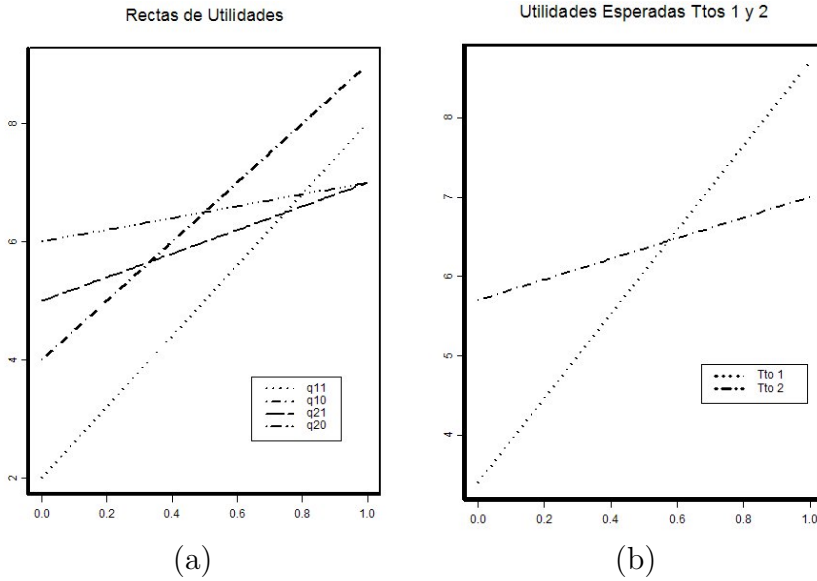


Figura 2.3: Comportamiento de conjuntos de utilidades

valores inferiores a $\alpha = 0.5$.

Para valores puntuales de α la utilidad esperada se calcula de igual manera que en el caso de una sola función de utilidad. Así que si se calcula en los extremos se puede obtener una representación gráfica de las utilidades esperadas de los dos tratamientos.

Por ejemplo tomando $\alpha = 0$, la utilidad esperada de aplicar el tratamiento 0 es:

$$\begin{aligned} U^*(0) &= \{P(X = 1)q_{01}^\alpha + P(X = 0)q_{00}^\alpha\} \\ &= \{P(X = 1)q_m(T_0, E_1) + P(X = 0)q_m(T_0, E_0)\}, \end{aligned}$$

y para el tratamiento 1,

$$\begin{aligned} U^*(1) &= \{P(X = 1)q_{11}^\alpha + P(X = 0)q_{10}^\alpha\} \\ &= \{P(X = 1)q_m(T_1, E_1) + P(X = 0)q_m(T_1, E_0)\}. \end{aligned}$$

Usar $\alpha = 1$ daría el valor máximo de la utilidad esperada para cada tratamiento. Por la linealidad de la esperanza, las utilidades esperadas para

todo $\alpha \in [0, 1]$ están incluidas en las rectas descritas por las utilidades esperadas de los valores extremos de α :

$$(1 - \alpha)U^*(0) + \alpha U^*(1).$$

Las rectas de utilidades esperadas de los tratamientos, se cortan en un punto que determina su igualdad. Gráficamente se observa cuál tratamiento prevalece sobre el otro, si lo hace sobre todo valor de utilidad posible o en qué momento un tratamiento pasa a ser el mejor.

En la Figura 2.3(b) se ve que la utilidad esperada del tratamiento 2 es superior cuando $\alpha < 0.57$. Esta información no es del todo útil porque no da respuesta concreta a la pregunta sobre el mejor tratamiento. Podrían considerarse para tomar la decisión elementos como la longitud del rango de la utilidad, o su pendiente. En este caso, se podría recomendar el tratamiento 2 por tener menor variación en su rango de utilidades esperadas y menor pendiente.

El método propuesto resulta útil cuando se observa la superioridad clara de uno de los tratamientos. Cuando se puede argumentar una elección a partir de la gráfica, es posible cometer errores generados por la subjetividad de la observación. Es necesario buscar un procedimiento matemáticamente justificado. Además este procedimiento no se puede implementar en un problema secuencial porque no es posible maximizar sumas de utilidades esperadas sumando los máximos de las utilidades esperadas.

Un método alternativo es mencionado por Rios Insua *et al.* (1997). Se propone descartar entre el conjunto de acciones posibles las que resultan con utilidades inferiores consistentemente. Para ello se define el concepto de acción no dominada.

$a \in A$ es *no dominada* si no existe otra alternativa $b \in A$ tal que $a < b$, es decir que la acción a sea menos preferible que la acción b , en términos de utilidades esperadas. Lo que se pide es que la utilidad esperada de la acción b no supere la utilidad esperada de la acción a bajo toda $u \in U$.

El método resultaría útil cuando se tiene un conjunto amplio de acciones posibles. Para el caso de elegir entre dos tratamientos, las posibles respuestas son que uno de los tratamientos domina al otro o que los dos son no dominados, es decir para algunas utilidades puntuales se prefiere el tratamiento 1 y para otras el tratamiento 2.

Un algoritmo que considera la clasificación de acciones no dominadas es propuesto y ejemplificado en el artículo de Christen, Müller, Wathen y Wolf (2004). Se realiza en este trabajo, la comparación de k tratamientos experimentales en un ensayo clínico fase II y un conjunto de posibles funciones de utilidad.

Capítulo 3

Utilidades en Medicina

3.1. Revisión Bibliográfica

En el análisis de costo-efectividad aplicado a problemas de decisión médica, interesa medir los beneficios de un tratamiento en términos de salud, no solo en términos económicos. (Parmigiani, 2002 p.78)

Existen diversos métodos para calcular beneficios en salud. Estos pueden clasificarse como psicométricos y de utilidad Lenert and Kaplan(2000). Los dos tipos de métodos consideran aspectos físicos, psicológicos, emocionales y de desempeño social, pero los métodos de utilidad incluyen la importancia dada por los individuos a un estado de salud, además son consecuentes con la teoría de utilidad lo que facilita su manejo en problemas de decisión.

Una fuente completa de estudios de costo-efectividad y costo-utilidad realizados entre 1976 al 2001, que incluye tablas de razones de costo-utilidad actualizadas al valor del dolar de 1998, se encuentra disponible en la página electrónica del *Harvard Center for Risk Analysis* <http://www.hsph.harvard.edu/cearegistry>. La información disponible permite comparaciones entre diferentes intervenciones a enfermedades diversas así como un reporte de los métodos usados en el análisis de cada estudio realizado.

Por otra parte, Velarde-Jurado y Avila-Figueroa (2002) del Instituto Nacional de Salud Pública de México, presentan la evolución del concepto de calidad de vida así como un inventario de 126 instrumentos utilizados para

medir calidad de vida en diversas poblaciones y enfermedades.

Diversos estudios de costo-efectividad han sido elaborados en el tema de las pruebas de diagnóstico de VIH. Tamashiro *et al.* (1993) presentan una serie de recomendaciones para reducir el costo de la aplicación de secuencias de pruebas en el diagnóstico del VIH y Bautista S., Dmytraczenko T., Kombre G. y Bertozzi S.(2003) realizaron un estudio sobre los costos de atención de VIH/SIDA en tres diferentes sistemas de salud de México.

Sobre las pruebas de monitoreo, el enfoque está en la forma de justificar el uso de pruebas de seguimiento más económicas. Jacobson *et al.* (2003) encontró en su estudio que para predecir conteos de CD4 inferiores a 350cel/ml un conteo Absoluto de Linfocitos de 1750cel/ml tenía una sensibilidad de 74 % y una especificidad del 73 %. Estos valores de sensibilidad y especificidad de la prueba mejoran para valores de CD4 más bajos.

3.1.1. QALYs: Quality Adjusted Life Years

Uno de los métodos para identificar beneficios en salud en la actualidad, son los QALYs que fueron introducidos por Fanshel y Bush (1970) como una alternativa para considerar las utilidades de las salidas en problemas de salud. Estos pertenecen al grupo de métodos de utilidad y se caracterizan por expresar estados de salud en términos de años de buena salud. Esto es, determinan las utilidades en términos del tiempo de permanencia en un estado inferior de salud equivalente a un periodo de tiempo en excelente salud. Para ello se construye una escala que asigna el 0 a la muerte y 1 a un estado de excelente salud; en el cálculo de los valores intermedios intervienen aspectos de tipo físico, de salud mental, desempeño social y laboral.

Un ejemplo sencillo para ilustrar la idea del funcionamiento de los QALYs es el siguiente. Si una persona vive 20 años bajo una condición de salud inferior a la excelente, y se dice que tiene una calidad de vida de 0.5, esto significa que requiere el doble del tiempo para vivir, o cumplir con un objetivo, lo que viviría si su salud fuera excelente. De esta manera los QALYs de su condición son 10, 10 años de vida ajustados por calidad.

Abellan y Pinto (2000) hacen un estudio de los QALYs como utilidades esperadas. En el artículo caracterizan los QALYs como funciones de utilidad

Von Newman-Morgenstern, para ello verifican las condiciones de mutua independencia entre QALYs, intercambio de tiempo proporcionalmente constante y neutralidad ante el riesgo. Introducen una nueva condición llamada de simetría que junto con la condición de independencia aditiva garantizan los demás supuestos de la caracterización de los QALYs.

En el artículo de Samsa *et al.* (1999), referenciado por Parmigiani (2002, p.62), se llevó a cabo un estudio para medir los QALYs de un *infarto mayor* en una población de riesgo creciente. El estudio consideró los candidatos a programas de prevención, algunos de ellos con problemas menores previos o enfermedades vasculares. Se definió como *infarto mayor* aquel que “deja un brazo, una pierna o un lado del cuerpo paralizado e inhabilita a la persona para cuidarse por si misma”. Además se asumió que “alguien después de sufrir un infarto de este tipo, permanecería en ese estado hasta la muerte.”

Se procedió a realizar a las personas del estudio preguntas de intercambios (trade-offs) como: “¿Preferiría vivir 10 años más después de un infarto mayor o vivir 8 años con excelente salud?” El tiempo propuesto por el entrevistador como de excelente salud fue variado, incluyendo decrementos de 3 y 6 meses, hasta alcanzar un punto de indiferencia, es decir una cantidad en años y meses que la persona estuviera dispuesta a considerar equivalente a vivir después de un infarto mayor. Entonces se dividió el valor de indiferencia entre 10. Si una persona eligió como punto de indiferencia, 7.25 años, sus QALYs estandarizados QS para un infarto mayor son de 0.725. Esta persona prefiere vivir 7.25 años de excelente salud a vivir 10 años después de sufrir un infarto mayor.

Para algunos autores, Baly (1999), los QS son conocidos como HRQL Health Related Quality of Life. Se obtiene de la asignación de valores entre 0 y 1 a los diferentes estados de salud, 0 a la muerte y 1 a la salud perfecta. Los QALYs se obtienen al multiplicar el HRQL por la duración en años de ese estado.

En cuanto a la estructura de los QALYs estandarizados, se consideran los siguientes supuestos:

- Un año de salud excelente equivale a 1.
- Un estado inferior de salud obtiene un valor inferior a 1.

- Dos años de vida con un estado de salud evaluado con $0.5QS$ equivalen a 1 año de salud excelente.
- Aplicar un tratamiento con un valor de $0.25QS$ a 4 individuos equivale a $1QS$.

3.1.2. Otras alternativas

Entre los métodos de cálculo del segundo tipo, sobresalen el Standard Gamble y los Time Trade Off. El primero usa probabilidades para determinar el punto en que resulta indiferente, a un individuo, un tratamiento que lo conduce con certeza a un estado de salud (intermedio entre el óptimo, excelente salud y el inferior, muerte), y un tratamiento que le proporciona una probabilidad de obtener el óptimo y su complemento el ínfimo. El punto de indiferencia es la utilidad del tratamiento.

Bajo este método se plantean al paciente dos acciones: a y a' . La acción a lo conduce a un estado óptimo de salud con una probabilidad α y a la muerte con probabilidad $1 - \alpha$. La acción a es llamada *Standard Gamble*. La acción a' , por otra parte, conduce con certeza al paciente a un estado intermedio de salud, por ejemplo requerir de una silla de ruedas para moverse. La idea del método es modificar la probabilidad α hasta que el paciente considere equivalente tomar cualquiera de las acciones.

Si la persona considera que el resultado de una acción es peor que la muerte se intercambian las salidas de las acciones, es decir se toma como *Standard Gamble* la acción que conduce con probabilidad α a la salud óptima y con probabilidad $1 - \alpha$ al resultado menos deseado.

En cuanto a los métodos de Time Trade Off, se establece el punto de indiferencia determinando equivalencias entre tiempos de vida en el estado óptimo y un estado inferior de salud. Entre estos se encuentran los QALYs.

Existen métodos similares que realizan equivalencias entre número de personas ayudadas por diferentes métodos. Por ello reciben el nombre de Person Trade Off.

Algunos de los inconvenientes de los métodos se relacionan con la falta

de habilidad numérica o de razonamiento probabilístico de las personas. Por ejemplo, un juego con un premio alto pero con baja probabilidad de ganar se prefiere a otro con bajo premio y alta probabilidad de ganar. Esta característica es muy común en el método de Standard Gamble.

Sin embargo, también es posible que la persona evalúe otros elementos en su decisión. Siguiendo con el ejemplo anterior, no sólo le interesa el monto del premio a obtener en el juego sino la ganancia o pérdida que tiene al invertir su tiempo jugando. Además es posible que tenga una actitud de apego o aversión hacia el riesgo que guíe su decisión. Este problema se presenta en los métodos de escala de razón, dificultando la posibilidad de conocer la evaluación de una condición de salud en sí misma.

Otro de los factores que afectan el funcionamiento de los métodos se refiere a la aplicación metodológica de los encargados de su práctica. Por ejemplo, el proceso de información al paciente empleado en la explicación del método puede afectar su decisión; si se utiliza material audiovisual que presenta personas de determinada raza o edad el individuo puede pensar que hay una relación directa entre la condición de salud en cuestión y el tipo de raza o de edad.

También es posible incurrir en errores de captura de la información, de cálculos no reales o incluso debidos a que la encuesta no incluye aspectos importantes a tener en cuenta en la decisión.

3.2. ¿Por qué utilizar QALYs?

Los QALYs resultan útiles por varias razones:

- La mayoría de las decisiones médicas se refieren a condiciones crónicas de salud, por lo que se toman sobre cantidades que representan el tiempo invertido en un estado de salud.
- No involucran el cálculo de probabilidades que puede confundir al paciente ni la posibilidad de una muerte inmediata, resultando más accesible para los pacientes.

- Son consistentes con la teoría de utilidad porque garantizan independencia entre longitud y calidad de vida, permiten asegurar proporcionalidad entre años de vida y son medidas cardinales que permiten agregar preferencias individuales.
- Es una medida significativa, válida, repetible y relevante. Por significativa se entiende su característica de ser comprensible para el usuario. Es válida porque mide lo que se desea medir. Es repetible porque se puede calcular en forma repetida con el mismo o con diferentes individuos y es relevante porque se refiere a la población afectada por la política de salud considerada, McGregor (2003).

Aún considerando las ventajas que presenta el uso de QALYs es importante reconocer, que como cualquier medida de utilidad, debe usarse con cautela. Específicamente, respecto al último punto mencionado en las ventajas, se encuentran estudios reales que corroboran las limitaciones y el mal uso de los QALYs. En el artículo de McGregor (2003) se menciona un estudio de costo-efectividad sobre la droga Sildenafil, empleada para corregir la disfunción erectil. Las conclusiones del estudio se basan en cálculos de dólares gastados por QALY obtenido. El estudio favorece el beneficio logrado a menor costo con Sildenafil, al compararlo con tratamientos referentes a problemas renales, de colesterol y diálisis. De esta manera se sugiere que es mejor invertir en el tratamiento con Sildenafil que en tratamientos para problemas completamente diferentes simplemente por ser más costo-efectivo. Un estudio comparativo de QALYs ganados debe considerarse en un contexto más amplio y servir como herramienta más que como elemento definitivo de decisión. Un error de interpretación en este sentido podría causar faltas graves en la toma de decisiones en salud pública.

En el mismo artículo se hace una crítica a la significancia del uso de los QALYs en este estudio. El medicamento realmente no causa un aumento significativo en la duración de la vida y no es clara la relación entre el aumento de calidad de vida en términos de cantidad de vida.

El supuesto de repetibilidad de cálculo de los QALYs, es difícil de asumir. Aún cuando la misma persona aplique en diferentes ocasiones los mismos cuestionarios para obtener los QALYs de un paciente respecto a un tratamiento específico no se garantizan salidas similares que hagan suponer

la repetibilidad de la medida.

Se discute la falta de relevancia refiriéndose nuevamente al estudio del Sildenafil. La respuesta sobre la importancia del uso del medicamento varía de acuerdo a la población consultada. Por ejemplo los hombres, sus esposas y médicos otorgan QALYs muy diferentes al tiempo de vida que cambiarían por evitar la impotencia. Un estudio al respecto podría variar dependiendo de la población objeto, el problema está en decidir cuál es la población que debería usarse en el problema de decisión.

Finalmente se menciona que a pesar de las dificultades, mientras no se encuentre una medida universal que posea todos los requerimientos necesarios, es importante contar con herramientas que permitan a los tomadores de decisión comparar las salidas de las acciones a considerar.

La descripción anterior se refiere a las políticas públicas empleadas para elegir la inversión en el tratamiento de una enfermedad A vs. el tratamiento para una enfermedad B, pero este punto no es de relevancia directa al asunto que interesa en este trabajo. Se menciona simplemente como parte de las críticas que existen en la literatura al uso de los QALYs.

3.3. Cálculo de utilidades usando QALYs

Frente a un problema de salud, un individuo puede tener alternativas de tratamiento que lo conduzcan, bajo diferentes probabilidades, a situaciones o consecuencias diferentes. La teoría de decisión puede guiar al individuo en su elección usando el cálculo de utilidades. Estas últimas consideran la probabilidad de las consecuencias determinadas por una acción tomada y les asignan un valor de utilidad con base en las preferencias del individuo.

La estimación de la utilidad de las consecuencias, en el caso de los QALYs, se hace determinando la equivalencia entre años de excelente salud y años con el estado de salud provocado por la enfermedad. Así, para una persona, diez años de su vida sometido a terapias antirretrovirales pueden equivaler a seis años de salud óptima, es decir, la utilidad de la consecuencia de permanecer siguiendo las terapias para esta persona es de 0.6 QALYs. Supóngase que esta persona tiene la posibilidad de decidir entre dos tratamientos. El primero

otorga un promedio de vida de 10 años con una probabilidad de no efectividad de 0.08 y el segundo tiene un promedio de 5 años de vida con una probabilidad de no efectividad de 0.01.

El cálculo de la utilidad esperada para las acciones a tomar el primer tratamiento y a' tomar el segundo tratamiento, será entonces:

$$U^*(a) = 0 * 0.08 + 0.6 * 10 * 0.92 = 5.52, \quad y$$

$$U^*(a') = 0 * 0.01 + 0.6 * 5 * 0.99 = 2.97.$$

Este resultado permite que la persona tenga un elemento de comparación entre las utilidades que le reportan las posibilidades de someterse a cada uno de los tratamientos. En este caso se favorece la opción 1, que otorga 5.52 años de vida ajustados por calidad, es decir 0.552 QALYs estandarizados.

En el problema de interés, la probabilidad de tener la condición E viene dada por las condiciones operativas de las pruebas y los resultados de sus aplicaciones y se calcula la utilidad de la secuencia de aplicaciones de las pruebas utilizando los QALYs de los tratamientos, como se explicó en la Sección 2.2.

En determinadas situaciones es posible no obtener un valor preciso de la utilidad que una consecuencia implica para una persona. (La persona puede estar muy afectada por su problema de salud o puede estar cansada de responder a los cuestionamientos sobre sus preferencias, etc.) Es posible sin embargo, obtener un rango de valores, es decir una cantidad de tiempo mínima y una máxima que una persona estaría dispuesta a considerar como equivalente a cierto tiempo de excelente salud. Esos valores pueden representarse como un conjunto de utilidades continuo. Un ejemplo de esta situación y la propuesta de solución se realizó en la Sección 2.2.

3.4. QALYs vs. costos

El objetivo de los análisis de costo-efectividad es establecer si una intervención proporciona beneficio a un costo aceptable (Parmigianni, 2000 p.78). En medicina este tipo de análisis es muy común en la comparación

de tratamientos, evaluación de proyectos de reforma sectorial y evaluación de presupuestos para incluir nuevas tecnologías. Usualmente se realiza un cociente entre los costos y los beneficios o efectividad.

En Parmigianni (2002 p.78) se menciona la propuesta de Pliskin respecto al cálculo de dolares por QALYs. Una segunda propuesta se refiere al cociente entre el costo esperado de una acción y la efectividad de la misma, en términos de QALYs.

Se considera z una salida o consecuencia particular, $e(z)$ la efectividad de una acción específica y $c_a(z)$ el costo de z obtenido bajo la acción a . El elemento de efectividad se relaciona con la utilidad por lo que se propone medirlo en QALYs y la parte del costo incluye el costo del procedimiento y aspectos causados por la acción como la prolongación de la vida o una discapacidad.

La efectividad esperada es:

$$E[a] = \int_Z \int_{\theta} e(z) p_a(z|\theta) p(\theta) d\theta dz, \quad (3.1)$$

que corresponde a la definición de utilidad esperada de una acción a al emplear $e(z)$ como $u(z)$.

El costo esperado es:

$$C[a] = \int_Z \int_{\theta} c_a(z) p_a(z|\theta) p(\theta) d\theta dz. \quad (3.2)$$

La razón promedio de costo-efectividad $C(a)/E(a)$ se compara con la obtenida bajo el cuidado estándar y se toma la decisión de acuerdo a la relación obtenida. Si el cociente de la acción a supera al obtenido bajo el cuidado estándar se preferirá esta nueva acción. Un ejemplo de su uso puede verse en el artículo de Ortún, V., Meneu, R. y Peiró S. (2004) sobre la evaluación del impacto de servicios sanitarios en España.

También se calcula la razón marginal de costo-efectividad para la acción a relativa a la acción a' de la siguiente manera:

$$\frac{C[a'] - C[a]}{E[a'] - E[a]}. \quad (3.3)$$

En un conjunto de acciones, las relaciones mencionadas se pueden usar para elegir la mejor, esto es, hacer comparaciones sucesivas entre parejas de acciones y eliminar las acciones que presentan menor valor de cociente. Secuencialmente se podría pensar en determinar la menor acción en cada etapa y en las etapas sucesivas partir de la mejor acción previa.

Otra posibilidad es calcular cocientes de efectividad por unidad de dinero invertido. En el contexto de este trabajo, la efectividad se mide en QALYs y el costo en dólares, así un cociente de estas cantidades indicará los QALYs ganados por dólar invertido resultando simplemente en el inverso del cociente de costo-efectividad propuesto anteriormente, como se observa enseguida:

$$\frac{E}{C} = \frac{q}{US\$}. \quad (3.4)$$

La dificultad principal de los análisis de costo-efectividad consiste en la agregación de valores individuales, es decir, en cómo tomar decisiones que afectan a toda una población con diferentes subgrupos determinados por las diferencias entre costos y las utilidades otorgados por ellos.

Un análisis de costo-efectividad referido al SIDA contempla varios factores: el costo de la intervención, el número de personas que alcanza, la eficacia de la intervención para cambiar conductas, además de las pérdidas morales y de productividad posible, entre otros (UCSF, 2002).

En el artículo de UCSF(2002) se mencionan resultados referentes a la preferencia de campañas dirigidas a poblaciones susceptibles como jóvenes, drogadictos y trabajadores sexuales antes que al público en general, el nivel de gasto ocasionado por el tratamiento anual de una persona afectada, algunos programas costo efectivos como el *Safer Choice* dirigido a escolares para promover el uso del condón y *Mpowerment* dirigida a hombres jóvenes homosexuales y bisexuales. Finalmente, en este artículo, se hace una recomendación de lo que falta por hacer en estudios de este tipo.

Por otra parte, Freedberg *et al.* (1998) determinan el impacto de diferentes estrategias para prevenir infecciones oportunistas en pacientes de VIH avanzado empleando un modelo de simulación de Markov y tomando mediciones de la calidad de vida obtenida en QALYs. Los resultados de esta investigación son posteriormente aplicados a un estudio de costo-efectividad

Freedberg *et al.* (2001) sobre la combinación de tratamientos ARV en la atención de pacientes con VIH.

Capítulo 4

Solución del problema

4.1. Cálculo de probabilidades

Para la solución del problema inicial: determinar la secuencia de pruebas óptima en N aplicaciones de T pruebas posibles para determinar si un paciente tiene o no la condición E y el tratamiento que debería seguir, se requiere la obtención de algunas probabilidades. Para ello es necesario recordar la notación a emplear:

- $P(X = 1)$ probabilidad a priori de que el paciente en cuestión posea la condición E de interés.
- c_t costo de aplicar la prueba t con $t = 1, 2, \dots, T$.
- q_{jk} QALY de aplicar el tratamiento j cuando la situación real de la persona es k , con $k = 1$ cuando la persona tiene la condición E y $k = 0$ en caso contrario.
- N número máximo de aplicaciones de pruebas a utilizar.
- Y_i^t el resultado de la prueba t aplicada en el intento i , que puede tomar valores 0 ó 1, para representar una prueba negativa o positiva para la condición E .
- $P(Y^t = 1|X = 0)$ probabilidad de falsos positivos de la prueba t
- $P(Y^t = 0|X = 1)$ probabilidad de falsos negativos de la prueba t .

Sea $P(X = 1|Y_1^{t_1} = y_1, Y_2^{t_2} = y_2, Y_3^{t_3} = y_3, \dots, Y_n^{t_n} = y_n)$ la probabilidad de que el paciente tenga la condición dados los resultados obtenidos en una serie de n aplicaciones de las pruebas, t_i corresponde a la prueba aplicada en el intento i de las T pruebas posibles.

Se conocen las probabilidades: $P(X = 1) = p$ probabilidad *a priori* de que el paciente tenga la condición de interés o prevalencia. $P(Y^t = 1|X = 0) = p_t$ probabilidad de obtener falsos positivos en la prueba t y $P(Y^t = 0|X = 1) = p'_t$ probabilidad de obtener falsos negativos en la prueba t .

Por el Teorema de Bayes, la probabilidad de tener la condición dado que la prueba t indica su presencia es:

$$P(X = 1|Y^t = 1) = \frac{P(Y^t = 1|X = 1)P(X = 1)}{P(Y^t = 1|X = 1)P(X = 1) + P(Y^t = 1|X = 0)P(X = 0)} = \frac{(1 - p'_t)p}{(1 - p'_t)p + p_t(1 - p)}. \quad (4.1)$$

La probabilidad de tener la condición dados los resultados de n_t aplicaciones de la prueba t , debe ser:

$$P(X = 1|Y_1^t = y_1, \dots, Y_n^t = y_n) = \frac{P(Y_1^t = y_1, \dots, Y_n^t = y_n|X = 1)P(X = 1)}{\sum_{x=0,1} P(Y_1^t = y_1, Y_2^t = y_2, \dots, Y_n^t = y_n|X = x)P(X = x)} = \frac{P(X = 1) \prod_{i=1}^{n_t} P(Y_i^t = y_i^t|X = 1)}{\sum_{x=0,1} \prod_{i=1}^{n_t} P(Y_i^t = y_i^t|X = x)P(X = x)} = \frac{(1 - p'_t)^{s_t} (p'_t)^{n_t - s_t} p}{(1 - p'_t)^{s_t} (p'_t)^{n_t - s_t} p + p_t^{s_t} (1 - p_t)^{n_t - s_t} (1 - p)}, \quad (4.2)$$

donde n_t se refiere al total de pruebas t aplicadas y s_t al total de pruebas t que resultaron positivas, esto es $s_t = \sum_{i=1}^{n_t} Y_i^t$. El resultado corresponde a la probabilidad de que una persona tenga la condición E dado el número de veces que la prueba t salió positiva de las n_t veces que se aplicó. Aquí se está asumiendo independencia entre los resultados de las pruebas condicionada a X lo cual parece razonable.

		Resultados de la Prueba	
		+	-
Realidad	+	True Positive	False Negative
	-	False Positive	True Negative

Tabla 4.1: Tabla de conteos

Las estadísticas n_t y s_t para las T pruebas son suficientes para X de tal manera que se puede considerar $D = (n_1, s_1, \dots, n_T, s_T)$ como el conjunto de datos necesario para determinar las probabilidades requeridas.

Si se consideran todas las aplicaciones de las T pruebas:

$$\begin{aligned}
 P(X = 1 | Y_1^{t1} = y_1^{t1}, Y_2^{t2} = y_2^{t2}, \dots, Y_n^T = y_n^T) = \\
 \frac{P(X = 1) \prod_{i=1}^n P(Y_1^{ti} = y_1^{ti} | X = 1)}{\sum_{x=0,1} P(Y_1^{ti} = y_1^{ti} | X = x) P(X = x)} = \quad (4.3) \\
 [p(1 - p'_1)^{s_1} (p'_1)^{n_1 - s_1} (1 - p'_2)^{s_2} (p'_2)^{n_2 - s_2} \dots (1 - p'_T)^{s_T} (p'_T)^{n_T - s_T}] * \\
 [p(1 - p'_1)^{s_1} (p'_1)^{n_1 - s_1} \dots (1 - p'_T)^{s_T} (p'_T)^{n_T - s_T} + \\
 (1 - p)(p_1)^{s_1} (1 - p_1)^{n_1 - s_1} \dots (p_T)^{s_T} (1 - p_T)^{n_T - s_T}]^{-1}
 \end{aligned}$$

Corresponde a la probabilidad $P(X = 1 | D)$, es decir la probabilidad de tener la condición dadas las salidas de todas las pruebas realizadas.

También es de interés calcular la probabilidad predictiva de que la prueba t resulte positiva dada la información disponible en un momento dado. Esto es:

$$P(Y^t = 1 | D) = \sum_{x=0,1} P(Y^t = 1 | X = x) P(X = x | D). \quad (4.4)$$

Usualmente la información requerida como probabilidades de falsos positivos y falsos negativos (Parmigianni 2002 p.7) viene dada en forma de conteos derivados de una serie de experimentos. Típicamente los experimentos consisten en clasificar muestras aleatorias de pacientes, de los que se conoce su situación real, y verificar la coincidencia y error de positivos y negativos de clasificación de la prueba, ver Tabla 4.1. Las entradas de la tabla definen los factores que determinan el rendimiento de una prueba.

La *sensibilidad* que se refiere a la capacidad de la prueba t para detectar correctamente la condición de interés en una persona, correspondiente a $P(Y^t = 1|X = 1)$ y la *especificidad* que determina el nivel de la prueba para reconocer una persona que no tiene la condición, esto es $P(Y^t = 0|X = 0)$. Una alta sensibilidad se relaciona con bajos resultados falso negativos y una especificidad alta con pocos resultados falso positivos.

Bajo el supuesto de un tamaño de muestra adecuado los conteos de la tabla aproximan las probabilidades mencionadas, luego la sensibilidad S y especificidad E de una prueba se pueden calcular de la siguiente manera:

$$S = \frac{TP}{TP + FN}, \quad y$$

$$E = \frac{TN}{TN + FP}.$$

El algoritmo empleado en la solución del árbol de decisión correspondiente al problema, requiere específicamente del cálculo de las probabilidades de *Falsos Negativos PFN* y *Falsos Positivos PFP* para cada una de las pruebas. Como la sensibilidad y especificidad corresponden a las probabilidades de *Verdaderos Positivos* y *Verdaderos negativos*, se puede calcular:

$$P(Y^t = 0|X = 1) = p'_t = 1 - P(Y^t = 1|X = 1) = 1 - S, \quad y \quad (4.5)$$

$$P(Y^t = 1|X = 0) = p_t = 1 - P(Y^t = 0|X = 0) = 1 - E.$$

Se calculan las probabilidades requeridas en términos de S y E porque usualmente las características operativas de las pruebas vienen dadas de esa manera. Por tanto se asumirá que los valores anteriores corresponden a las probabilidades buscadas.

4.2. Solución al problema con una sola función de utilidad

Como se mencionó en el Capítulo 2, el interés al resolver el árbol de decisión que representa el problema es saber si conviene continuar con la aplicación de pruebas y cuál prueba en ese caso ó detener su aplicación y

4.2. SOLUCIÓN AL PROBLEMA CON UNA SOLA FUNCIÓN DE UTILIDAD 37

	$X = 0$	$X = 1$
Tto 0	q_{00}	q_{01}
Tto 1	q_{10}	q_{11}

Tabla 4.2: Tabla de conjuntos de QALYs

optar por recomendar uno de los tratamientos. Para ello es necesario calcular la utilidad esperada de parar, es decir el QALY esperado de parar y a la vez considerar el costo de tomar esa decisión.

Los costos de las pruebas están dados por c_1, c_2, \dots, c_T y el esquema de utilidades, en el caso de una sólo función de utilidad, viene dado por la Tabla 4.2. El problema de decisión se puede plantear ahora considerando las acciones: *Seguir* con el número de pruebas posibles de las cuáles se elegirá la que reporte una utilidad máxima y *Parar* al cabo de la cual se debe recomendar uno de los tratamientos.

Como se explicó en la Sección 2.2, es necesario calcular las utilidades esperadas para cada nodo de decisión del árbol. Si se tiene como límite realizar N aplicaciones se tienen T^N posibilidades de secuencias de pruebas. La información de una secuencia específica de pruebas se recoge en $D = (n_1, s_1, \dots, n_T, s_T)'$.

La utilidad de parar, representada en un nodo de decisión, corresponde a un máximo, como se explicó en la Sección 2.2. En general, es el máximo de los QALYs esperados de cada tratamiento dada la información obtenida hasta el momento. Esto es:

$$q_s^* = \max\{P(X = 1|D)q_{01} + P(X = 0|D)q_{00}, P(X = 1|D)q_{11} + P(X = 0|D)q_{10}\}.$$

Los QALYs de los nodos aleatorios se calculan como utilidades esperadas, teniendo en cuenta las salidas posibles de las diferentes pruebas. Es decir utilizando las probabilidades predictivas de cada una de las pruebas dada la información disponible en D:

$$P(Y^t = y|D) = \sum_{x=0,1} P(Y^t = 1|X = x)P(X = x|D).$$

Siguiendo el procedimiento de *Backward Induction* se calculan las respectivas utilidades de los nodos del árbol y se va definiendo en cada momento el paso a seguir dada la probabilidad de tener la condición de interés y una función de utilidad específica.

Además, para incluir los costos en la decisión se define la siguiente estrategia. Se tomará como costo inicial el valor C que puede corresponder al máximo de los costos de los tratamientos. Este costo se refiere también al máximo que debería pagarse por iniciar un tratamiento sin información proveniente de las pruebas. Si se decide realizar pruebas, el costo incrementa. Así $C + \sum_{i=1}^n c_{t_i}$ será el nuevo costo después de realizar las n pruebas t_1, t_2, \dots, t_n .

Si en un punto específico del árbol, por ejemplo al iniciar el árbol se quiere comparar la ventaja de no realizar pruebas o empezar con la prueba 1 ó 2, se tiene C y $C_1 = C + c_1$ costo que incluye el costo de aplicar el tratamiento y el costo de realizar la prueba 1 en primer lugar y $C_2 = C + c_2$ en forma equivalente para iniciar con la prueba 2. Además la utilidad de parar está dada por q_s^* , la de continuar con la prueba 1 por q_1^* y respectivamente q_2^* para continuar con la segunda prueba. Con ésta información se comparan los cocientes de costo efectividad:

$$\frac{q_s^*}{C} \quad vs. \quad \frac{q_1^*}{C_1}, \quad y \quad (4.6)$$

$$\frac{q_s^*}{C} \quad vs. \quad \frac{q_2^*}{C_2}.$$

Este valor se interpreta como la cantidad de años de buena salud ganados por cada dólar invertido en el tratamiento. El procedimiento completo empleando *Backward Induction* para definir la secuencia óptima de acciones se implementó en el programa elaborado en C++. Ejemplos de su aplicación se mostrarán en la Sección 4.4.

4.2. SOLUCIÓN AL PROBLEMA CON UNA SOLA FUNCIÓN DE UTILIDAD 39

	$X = 0$	$X = 1$
Tto 0	$q_{00} = 10$	$q_{01} = 2$
Tto 1	$q_{10} = 1$	$q_{11} = 7$

Tabla 4.3: Función de utilidad para el ejemplo de aplicación del algoritmo

4.2.1. Ejemplo de la aplicación del algoritmo de *Backward Induction*

Para fijar ideas, a continuación se muestran los cálculos realizados para resolver el caso $T = 2$ y $N = 2$, 2 pruebas con un límite de 2 aplicaciones. Se supondrá que la condición a evaluar es la presencia o ausencia de una enfermedad. Se utilizaron las probabilidades: $p = P(X = 1) = 0.05$ probabilidad a priori de tener la enfermedad, $p'_1 = P(Y^1 = 0|X = 1) = 0.038$ y $p_1 = P(Y^1 = 1|X = 0) = 0.02$ probabilidades de falsos negativos y falsos positivos de la prueba 1 y $p'_2 = P(Y^2 = 0|X = 1) = 0.001$ y $p_2 = P(Y^2 = 1|X = 0) = 0.012$ probabilidades de falsos negativos y falsos positivos de la prueba 2. Se empleó un costo de tratamiento de $C = US\$50$ y los costos de las pruebas son de $c_1 = US\$1$ y $c_2 = US\$11$ respectivamente. Finalmente la función de utilidad empleada es la descrita por la Tabla 4.3.

El árbol de la Figura 4.1 representa todas las acciones posibles a seguir en este caso. Los números del 1 al 36 denotan las utilidades en cada nodo del árbol. El procedimiento de *Backward Induction* indica que debe iniciarse por los extremos del árbol, es decir con las utilidades de los nodos 1 a 16. El cálculo de estas utilidades requiere maximizar la probabilidad de tener o no la enfermedad dado el resultado de las pruebas empleadas en las dos primeras aplicaciones. La utilidad resultante corresponde al mayor cociente del QALY esperado con la información dada y el costo total que incluye el costo del tratamiento y de las pruebas realizadas. A continuación se presenta el cálculo completo del nodo 1.

Primero se calcula la probabilidad, $P(X = 1|Y_2^1 = 1, Y_1^1 = 1)$ correspondiente a tener la enfermedad dado que las dos aplicaciones que se hicieron de la prueba 1 resultaron positivas. Evaluando la expresión (4.3)

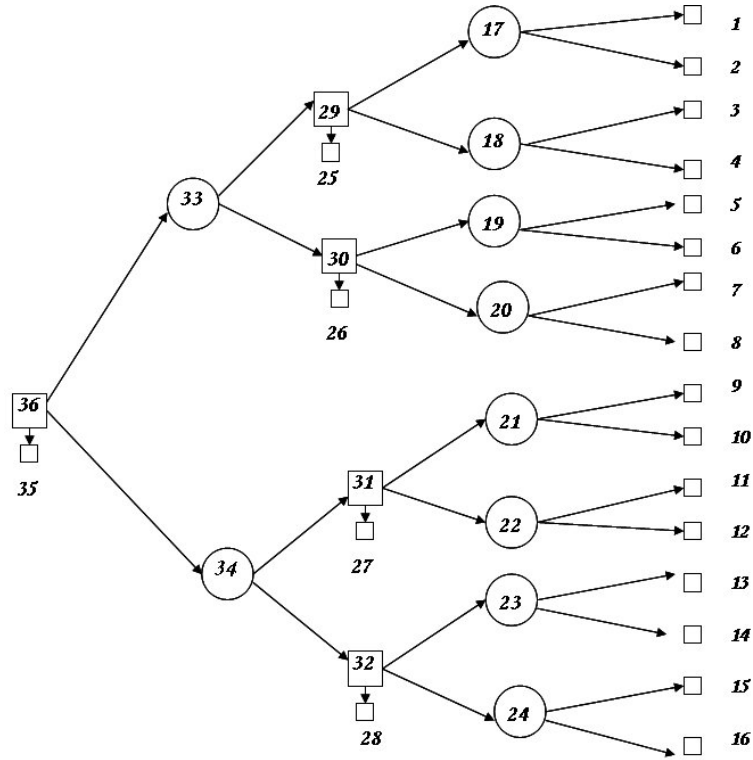


Figura 4.1: Árbol de decisión para el ejemplo de aplicación del algoritmo

para los valores $n_1 = 2$, $s_1 = 2$ se obtiene:

$$P(X = 1 | Y_1^1 = 1, Y_2^1 = 1) = \quad (4.7)$$

$$\frac{(1 - p'_1)^{s_1} (p'_1)^{n_1 - s_1} p}{(1 - p'_1)^{s_1} (p'_1)^{n_1 - s_1} p + (p_1)^{s_1} (1 - p_1)^{n_1 - s_1} (1 - p)} =$$

$$\frac{(1 - p'_1)^2 p}{(1 - p'_1)^2 p + (p_1)^2 (1 - p)} =$$

$$\frac{(1 - 0.038)^2 (0.05)}{(1 - 0.038)^2 (0.05) + 0.02^2 (0.95)} =$$

$$0.991854.$$

Ahora usando la probabilidad de la expresión (4.7) se obtiene la utilidad

4.2. SOLUCIÓN AL PROBLEMA CON UNA SOLA FUNCIÓN DE UTILIDAD 41

del nodo 1, como el máximo de los cocientes:

$$\{P(X = 1|Y_2^1 = 1, Y_1^1 = 1) * q_{01} + (1 - P(X = 1|Y_2^1 = 1, Y_1^1 = 1)) * q_{00}\} / (C + 2c_1)$$

y

$$\{P(X = 1|Y_2^1 = 1, Y_1^1 = 1) * q_{11} + (1 - P(X = 1|Y_2^1 = 1, Y_1^1 = 1)) * q_{10}\} / (C + 2c_1).$$

Donde $C + 2c_1$ representa el costo de aplicar tratamiento más el costo de aplicar dos veces la prueba 1 para diagnosticar la enfermedad. Debe notarse que la primera expresión corresponde al caso de aplicar el tratamiento 0 y la segunda a la aplicación del tratamiento 1. Finalmente se asigna al nodo 1 del árbol el resultado máximo de los cocientes, esto es: $\max\{0.03971467, 0.13367553\}$ es decir 0.13367553, que implica suponer positivo al paciente y elegir el tratamiento 1.

En el Anexo A en la Figura A.2 se muestran los cálculos correspondientes a cada uno de los nodos del árbol además de la salida del programa en la Figura A.1. Para continuar la explicación se tomarán los resultados de los nodos 1 a 4. Con los resultados de los nodos 1 y 2 se calcula el valor del nodo 17, que corresponde a una utilidad esperada:

$$P(Y_2^1 = 1|Y_1^1 = 1) * \text{nodo1} + P(Y_2^1 = 0|Y_1^1 = 1) * \text{nodo2}.$$

El cálculo de la probabilidad involucrada en esta utilidad es:

$$P(Y_2^1 = 1|Y_1^1 = 1) = \tag{4.8}$$

$$\begin{aligned} P(Y_2^1 = 1|X = 1)P(X = 1|Y_1^1 = 1) + P(Y_2^1 = 1|X = 0)P(X = 0|Y_1^1 = 1) = \\ (1 - p'_1) \frac{(1 - p'_1)p}{(1 - p'_1)p + (p_1)(1 - p)} + (p_1) \frac{p_1(1 - p)}{(1 - p'_1)p + p_1(1 - p)} = \\ (1 - 0.038) \frac{(1 - 0.038)(0.05)}{(1 - 0.038)(0.05) + (0.02)(0.95)} + \\ 0.02 \frac{0.02(1 - 0.05)}{(1 - 0.038)(0.05) + (0.02)(1 - 0.05)} = \\ 0.7039374. \end{aligned}$$

Luego el valor del nodo 17, es:

$$\begin{aligned} P(Y_2^1 = 1|Y_1^1 = 1) * nodo1 + P(Y_2^1 = 0|Y_1^1 = 1) * nodo2 = & \quad (4.9) \\ 0.7039374 * 0.13367553 + (1 - 0.7039374) * 0.17855560 = & \\ & 0.1469628. \end{aligned}$$

En forma equivalente se calculan los nodos 18 a 24. Los nodos 25 a 28 corresponden a máximos de los cocientes de utilidades esperadas bajo la probabilidad de tener la enfermedad dado el resultado de la primer prueba. El nodo 25, es de la forma:

$$\begin{aligned} max \left\{ \frac{P(X = 1|Y_1^1 = 1) * q_{01} + P(X = 0|Y_1^1 = 1) * q_{00}}{C + c_1}, \right. & \quad (4.10) \\ \left. \frac{P(X = 1|Y_1^1 = 1) * q_{10} + P(X = 0|Y_1^1 = 1) * q_{11}}{C + c_1} \right\}, \end{aligned}$$

donde $C + c_1$ corresponde al costo de aplicar un tratamiento más el costo de aplicar la prueba 1, una vez. La probabilidad $P(X = 1|Y_1^1 = 1)$ se calculó como parte de la expresión 4.8 y es 0.71684053, luego el máximo de los cocientes de 4.10 se evalúa como:

$$\begin{aligned} max \left\{ \frac{0.71684053 * 2 + (1 - 0.71684053) * 10}{50 + 1}, \right. & \quad (4.11) \\ \left. \frac{0.71684053 * 1 + (1 - 0.71684053) * 7}{50 + 1} \right\} = \\ max \{0.08363286, 0.103942\}. \end{aligned}$$

Esto implica que el nodo 25 vale 0.10394202 y recomienda seguir el tratamiento 1 si se decidiera detener el experimento. Los nodos 27 a 32 corresponden a los máximos de parar tras tener el resultado de la primera prueba y continuar con la segunda aplicación. Por ejemplo el nodo 29 es

$$max \{0.1469628, 0.1258010, 0.103942\},$$

es decir, vale 0.14696284 e indica como óptima la decisión de continuar el experimento aplicando la prueba 1 por segunda ocasión.

4.2. SOLUCIÓN AL PROBLEMA CON UNA SOLA FUNCIÓN DE UTILIDAD 43

Los nodos 33 y 34 se obtienen como utilidades esperadas dadas por:

$$P(Y_1^1 = 1) * nodo29 + P(Y_1^1 = 0) * nodo30$$

y

$$P(Y_1^2 = 1) * nodo31 + P(Y_1^2 = 0) * nodo32.$$

Para el nodo 33,

$$\begin{aligned} P(Y_1^1 = 1) &= P(Y_1^1 = 1|X = 1)P(X = 1) + P(Y_1^1 = 1|X = 0)P(X = 0) \quad (4.12) \\ &= (1 - p'_1)p + p_1(1 - p) \\ &= 0.0671. \end{aligned}$$

El nodo 33 entonces toma el valor 0.19248474 al resolver

$$P(Y_1^1 = 1) * nodo29 + P(Y_1^1 = 0) * nodo30 = 0.19248474. \quad (4.13)$$

El nodo 35 correspondiente a la acción de no realizar pruebas se calcula como:

$$\begin{aligned} \max \left\{ \frac{P(X = 1) * q_{01} + P(X = 0) * q_{00}}{C}, \frac{P(X = 1) * q_{10} + P(X = 0) * q_{11}}{C} \right\} \\ \max \left\{ \frac{0.05 * 2 + 0.95 * 10}{50}, \frac{0.05 * 1 + 0.95 * 7}{50} \right\} \\ \max \{0.192, 0.134\}. \quad (4.14) \end{aligned}$$

Luego, si se decidiera no realizar pruebas se recomendaría el tratamiento 0, con un valor de 0.192 de QALY ganado por dólar invertido. Finalmente el nodo 36 corresponde al máximo de las acciones, parar, empezar con la prueba 1 y empezar con la prueba 2. Esto es $\max \{0.192, 0.1924847, 0.16106725\}$, es decir vale 0.1924847, que indica empezar con la prueba 1.

Si ahora se hace el seguimiento de las mejores decisiones en cada etapa del árbol, se observará que la secuencia óptima, descrita en la Figura 4.2, sugiere empezar por la prueba 1. Si el resultado es negativo declarar negativa la enfermedad en la persona y recomendarle el tratamiento 0. Si la respuesta de la primera prueba es positiva se debe aplicar nuevamente la prueba 1 y seguir el tratamiento 0 si la respuesta es negativa ó el tratamiento 1 si la respuesta es positiva. Es de notar que en este caso, la prueba 2 simplemente no se

utiliza. La secuencia óptima obtenida en el programa puede verse en el archivo `output10.html` y sus datos de ingreso en el archivo `input10.txt` en la dirección <http://www.cimat.mx/~jac/material/CaroTesis/ejemplos.zip>. En esta página se encuentran todos los ejemplos del trabajo.

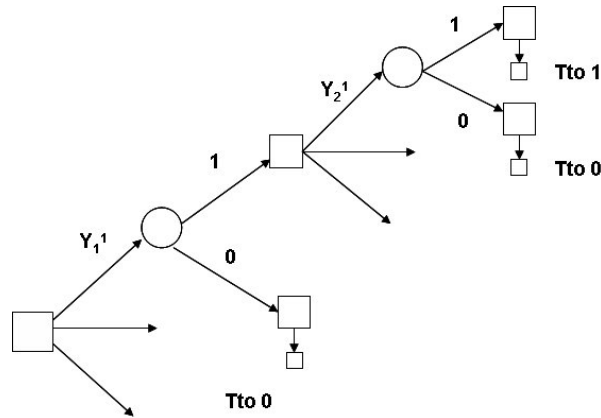


Figura 4.2: Secuencia óptima para el ejemplo de aplicación del algoritmo.

4.3. Ejemplos y Discusión

4.3.1. Ejemplo 1. Estrategia III UNAIDS/WHO

El primer ejemplo corresponde a las pruebas de diagnóstico de VIH. Se quiere comparar los resultados obtenidos con la propuesta de UNAIDS/WHO (1997) para la aplicación de pruebas de diagnóstico. En este ejemplo la condición a evaluar es la presencia del VIH en un paciente, $k = 1$ cuando el virus está presente y $k = 0$ en caso contrario. Los valores de probabilidades de falsos positivos y negativos, al igual que los costos, se obtuvieron del artículo de Miró *et al.* (2004) y del reporte de WHO (2004).

La función de utilidad se definió suponiendo que el tratamiento 0 es adecuado cuando se declara negativa la enfermedad y el tratamiento 1 cuando se afirma que el paciente posee la enfermedad. Específicamente se utilizó $q_{00} = 10$ para no tener la condición y aplicar el tratamiento 0, lo que representa los 10 años de vida saludable que tendría la persona al

someterse a ese tratamiento sin tener el virus y $q_{10} = 2$ para tener la condición para el mismo tratamiento, esto significa que la persona ganaría 2 años de buena salud por seguir el tratamiento cuando tiene el virus. Respectivamente $q_{10} = 1, q_{11} = 7$ para el tratamiento 1. Se utilizó un límite de dos pruebas, $N = 2$; una prevalencia de 0.05 esto es probabilidad a priori $P(X = 1) = 0.05$, y las probabilidades de falsos positivos y negativos calculados como $P(Y^1 = 1|X = 0) = 0.03, P(Y^1 = 0|X = 1) = 0.21$ para la prueba $t = 1$, prueba de ELISA; de $P(Y^2 = 1|X = 0) = 0.02$ y $P(Y^2 = 0|X = 1) = 0.038$ para $t = 2$, Pruebas Rápidas y $P(Y^3 = 1|X = 0) = 0.012, P(Y^3 = 0|X = 1) = 0.001$ para $t = 3$, prueba Western Blot. Los costos de las pruebas son $c_1 = US\$0.5, c_2 = US\$1, c_3 = US\$11$ y el costo inicial propuesto es de $C = US\$100$.

Se consideran todas las posibles salidas y posibles secuencias de pruebas cuando se pueden realizar sólo dos pruebas. El algoritmo establece como óptima la secuencia presentada en la Figura 4.3 con una utilidad de 0.097308 de QALYs obtenidos por dólar invertido. Esto es, iniciar aplicando la prueba 1, Y_1^1 , una prueba de ELISA; seguir con la prueba 2, Y_2^2 , si resulta positiva ó parar si el resultado es negativo. Los QALYs por dólar de realizar una segunda aplicación de la prueba 2 son de 0.0795226 y 0.0986057 en el caso de parar, sugiriendo el tratamiento 0 (Ver salida en el archivo output11.html). La salida completa para este caso se puede observar en el archivo output12.html, en la dirección electrónica mencionada, si se quieren verificar otras posibilidades de inicio de la secuencia.

El algoritmo de UNAIDS/WHO (1997) incluye tres estrategias de acción de acuerdo al objetivo de las pruebas y la prevalencia de la infección. En la Figura 4.4 se observa una clasificación de las estrategias y en la Figura 4.5 su representación secuencial. En el ejemplo 1, con una prevalencia de 0.05 y suponiendo que se desea diagnosticar a un paciente asintomático la recomendación es seguir la estrategia III. Esto es Y_1^1 , si $Y_1^1 = 0$ reportar negativo al paciente. Si $Y_1^1 = 1$ aplicar la prueba 2, Y_2^2 ; si $Y_2^2 = 0$ repetir las pruebas 1 y 2. Si estas son negativas declarar negativo al paciente, si $Y_3^1 = 1$ y $Y_4^2 = 0$ ó $Y_3^1 = 1$ y $Y_4^2 = 1$ aplicar la prueba 3, esto es Y_5^3 . Si $Y_2^2 = 1$ aplicar la prueba 3, Y_3^3 . Finalmente se declara positivo, sólo cuando las tres pruebas han resultado positivas, sin interesar el número de aplicaciones que se hayan efectuado. Los demás casos se consideran indeterminados. Esta secuencia se puede observar también en el árbol de la Figura 4.6.

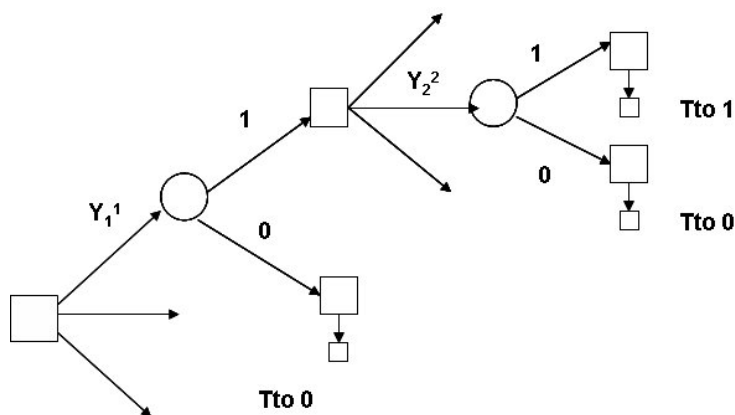


Figura 4.3: Secuencia Óptima para el ejemplo 1 de pruebas de diagnóstico.

Esta descripción no corresponde a la secuencia óptima propuesta bajo el modelo de Backward Induction. Para empezar asume un límite de 5 aplicaciones de pruebas. Si se resuelve el problema con $N = 5$, la secuencia óptima es la descrita por árbol de la Figura 4.7. Puede verse que se hace referencia continua al uso de las pruebas 1 y 2. Se inicia con la prueba 1 y si $Y_1^1 = 0$ se declara negativo al paciente, en caso contrario se aplica por segunda vez la prueba 1, Y_2^1 y si $Y_2^1 = 1$ nuevamente la prueba 1 pero si $Y_2^1 = 0$ se aplica la prueba 2, que confirma un paciente negativo en caso de resultar negativa. Si $Y_3^2 = 1$ se aplica Y_4^2 reportando positivo al paciente si $Y_4^2 = 1$ y sugiriendo aplicar nuevamente la prueba 2 en el quinto intento para definir la condición del paciente. (Ver salida del programa en output13.html)

Si se permite ahora la aplicación de 3 pruebas, $N = 3$, conservando los demás datos, el resultado es el mismo, se aplican sólo dos pruebas. Es decir se propone usar la prueba 1, Y_1^1 y si el resultado es positivo aplicar la prueba 2, Y_2^2 . En caso de un negativo en la primera prueba la recomendación es declarar negativo al paciente y aplicarle el tratamiento 0. (Ver archivo output14.html). Si se permiten 2 pruebas, $N = 2$ y se usa una prevalencia superior, $P(X = 1) = 0.20$ la secuencia óptima es otra vez la misma a la representada en el árbol de la Figura 4.3, aunque con diferentes utilidades. (Consultar output15.html)

WEEKLY EPIDEMIOLOGICAL RECORD, No. 12, 21 MARCH 1997 • RELLEVÉ ÉPIDÉMIOLOGIQUE HEBDOMADAIRE, No 12, 21 MARS 1997

Objective of testing Objectif du dépistage	Prevalences of infection Prévalences de l'infection	Testing strategy Stratégie de dépistage
<i>Transfusion/transplant safety</i> <i>Sécurité des transfusions/transplantations</i>	All prevalences Toutes prévalences	I
<i>Surveillance</i>	>10%	I
	≤10%	II
<i>Diagnosis – Diagnostic</i>	Clinical signs/symptoms of HIV infection ^a Signes cliniques/symptômes de l'infection à VIH ^a	>30% I
	≤30%	II
	Asymptomatic – Asymptomatique	>10% II
	≤10%	III

^a World Health Organization. *In terim proposal for a WHO staging system for HIV infection and disease* (WER No. 29, 1990, pp. 221-228). – Organisation mondiale de la Santé. *Echelle provisoire OMS proposée pour la détermination des stades de l'infection à VIH et de la maladie* (REH No 29, 1990, pp. 221-228).

Figura 4.4: Recomendaciones de aplicación de estrategias ONUSIDA

4.3.2. Ejemplo 2. Pruebas de Monitoreo de SIDA

Como se mencionó en la introducción, el monitoreo del SIDA se realiza a través de cuatro pruebas posibles. En este caso no existe en la literatura una estrategia definida de secuencia de pruebas a seguir. En algunos casos, se decide a partir de la combinación de resultados de las pruebas CD4 y Carga Viral. Estudios recientes intentan reconocer la validez de aplicar el Conteo Total de Linfocitos como prueba diagnóstica e identificar valores equivalentes a los de CD4 y Carga Viral para tomar decisiones y ahorrar en costos de aplicación (Jacobson *et al.* 2003).

Otros estudios procuran incluir el seguimiento del Índice de masa corporal IMC como indicador del avance de la enfermedad. La pérdida de peso involuntaria superior al 10 % aunada a síntomas como la persistencia de fiebre alta sugiere el avance de la enfermedad. Un avance en este sentido podría contribuir mucho a la toma de decisiones en países con grandes prevalencias de la enfermedad y escasos recursos (Díaz y Lubian, 1998).

En este ejemplo la condición *E* se refiere a la necesidad de iniciar un

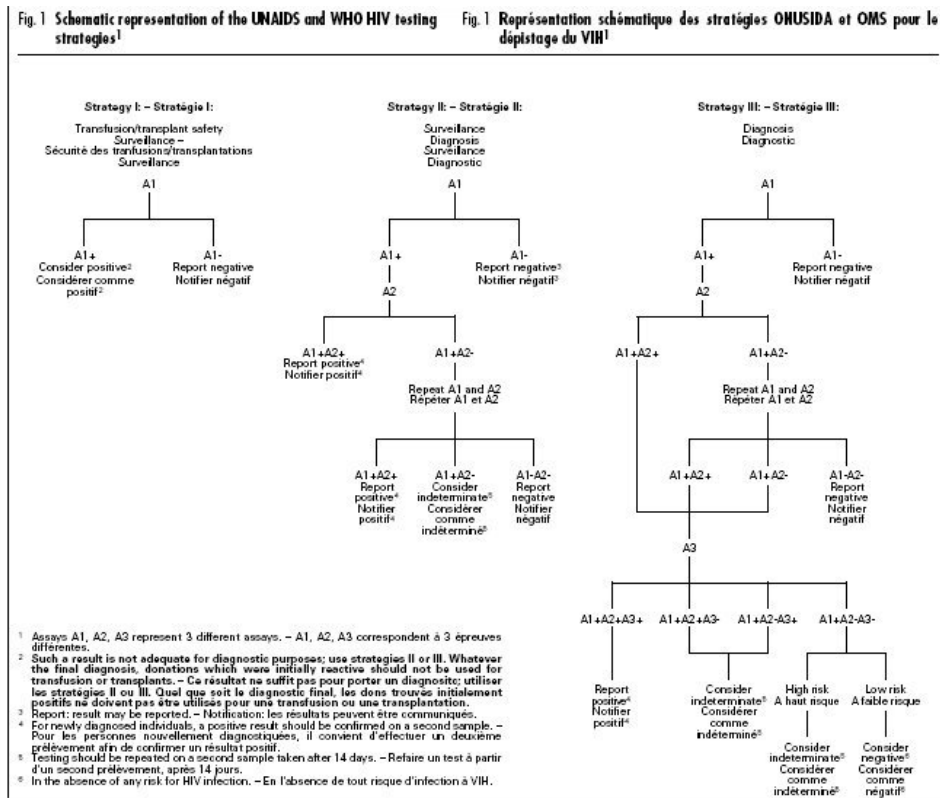


Figura 4.5: Representación de estrategias ONUSIDA

tratamiento antiretroviral o cambiar el que se viene aplicando. Esto se traduce en $k = 1$ si los niveles de defensa de la persona han disminuido, su carga viral ha aumentado ó ha perdido peso de manera involuntaria. Es bajo estos criterios que realmente se declara que la persona padece de SIDA. Por tanto se puede decir que $k = 1$ corresponde a tener SIDA. Las pruebas que se van a utilizar en este ejemplo son IMC, Linfocitos Totales, CD4 y Carga Viral. En este caso se requiere un supuesto adicional al problema referente a la aplicación consecutiva de pruebas, se asumirá aquí que la siguiente aplicación no es inmediata sino que se llevará a cabo cierto tiempo después, un mes por ejemplo. La frecuencia ideal de aplicación de las pruebas es un problema que no es de interés en el desarrollo del presente trabajo.

Se consideraron las probabilidades de falsos positivos y negativos de cada

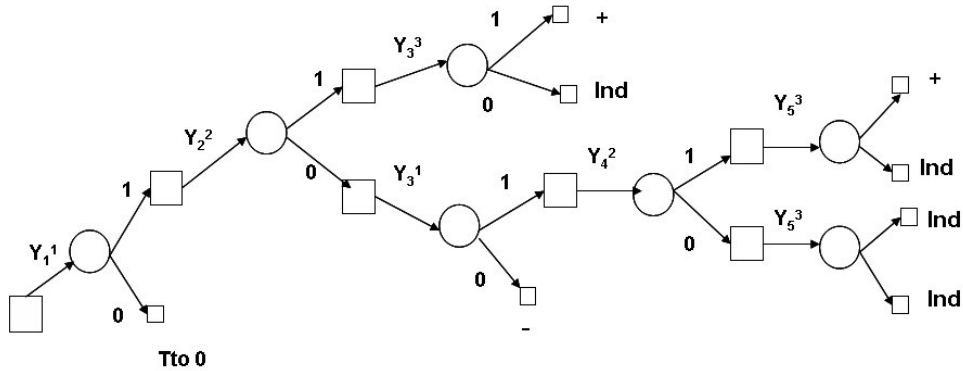


Figura 4.6: Árbol de decisiones para ONUSIDA

prueba como $P(Y^1 = 1|X = 0) = 0.18$, $P(Y^1 = 0|X = 1) = 0.25$ para t_1 IMC, $P(Y^2 = 1|X = 0) = 0.365$, $P(Y^2 = 0|X = 1) = 0.165$ para t_2 Linfocitos Totales, $P(Y^3 = 1|X = 0) = 0.043$, $P(Y^3 = 0|X = 1) = 0.194$ para t_3 CD4 y $P(Y^4 = 1|X = 0) = 0.05$, $P(Y^4 = 0|X = 1) = 0.01$ para t_4 Carga Viral. Estos valores se obtuvieron de UNAIDS/WHO (2001) y Díaz y Lubian (1998). La función de utilidad fue la misma empleada en el ejemplo 1; los costos de las pruebas son $c_1 = 0$, $c_2 = 2$, $c_3 = 40$, $c_4 = 70$ y el costo inicial de $c = 1000$ dólares.

Para la secuencia de resultados descrita por el árbol de la Figura 4.8, un límite de 5 pruebas $N = 5$ y usando una prevalencia de $P(x = 1) = 0.06$ se recomienda iniciar con la prueba 1. La prueba 1 es la más recomendada sólo se hace referencia a la prueba 4 en la última aplicación cuando no se ha logrado decidir previamente la condición de la persona. Por ejemplo si el primer resultado es negativo $Y_1^1 = 0$ se debe aplicar nuevamente la prueba 1 y si nuevamente resulta negativa, $Y_2^1 = 0$ declarar negativo al paciente. Si $Y_2^1 = 1$ se recomienda aplicar la prueba 1 nuevamente, se declara negativo al paciente si $Y_3^1 = 0$, si $Y_3^1 = 1$ se aplica nuevamente la prueba 1 y si $Y_4^1 = 0$, se declara negativo pero si no, se aplica la prueba 4 en la quinta aplicación. Esto es Y_5^4 de su resultado se obtiene una respuesta definitiva. Un resultado interesante es que se determina el positivo cuando las cuatro primeras aplicaciones de la prueba 1 son positivas y que no se usan las pruebas 2 y 3. Los cálculos correspondientes a la secuencia óptima se pueden observar en output21.html.

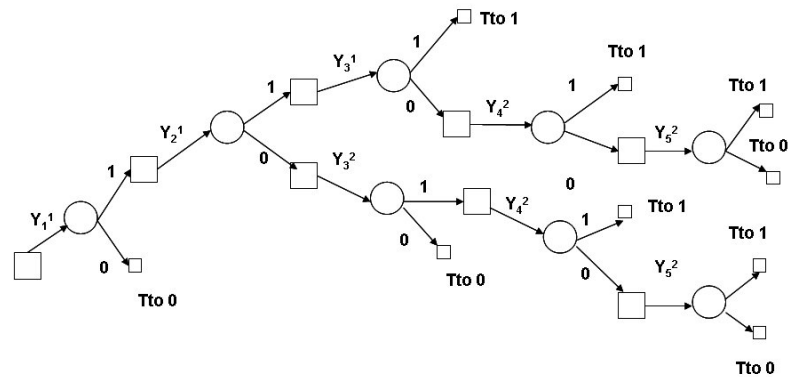


Figura 4.7: Secuencia óptima alternativa a la estrategia de ONUSIDA

También se puede notar que como $(n_1, s_1, \dots, n_n, s_n)$ es una estadística suficiente, los nodos con la misma estadística son iguales. Por ejemplo los nodos marcados con la misma letra en la Figura 4.8. Los nodos A tienen como estadística $(n_1 = 2, s_1 = 1)$, los nodos B al conjunto $(n_1 = 3, s_1 = 2)$ y los nodos C al conjunto $(n_1 = 4, s_1 = 3)$. Por tanto es posible reducir el árbol del ejemplo 2 al representado en la Figura 4.9. En resumen la prueba 1 se aplica máximo 4 veces, si resultan 2 negativas se declara al paciente negativo para SIDA y si las 4 pruebas del tipo 1 resulta positivas se declara positivo. En caso de que sólo una de las pruebas 1 sea negativa se debe aplicar la prueba 4 para decidir. Además las pruebas 2 y 3 no se utilizan.

En el mismo ejemplo, aplicar la prueba más costosa Y_1^4 la primera vez, implica declarar negativa inmediatamente a la persona y si el resultado es positivo se sugiere aplicar la prueba 1, si $Y_2^1 = 1$ nuevamente la prueba 1 y si $Y_3^1 = 1$ otra vez la prueba 1 y finalmente en su cuarta aplicación y resultado positivo $Y_4^1 = 1$ se reporta positivo al paciente. La secuencia completa cuando se decide empezar por la prueba 4 se representa en el árbol de la Figura 4.10 y su salida puede verse en [output22.html](#).

Emplear una prevalencia alta 0.35 y la posibilidad de utilizar únicamente dos pruebas sugiere iniciar con la prueba 1, Y_1^1 . Si el resultado es positivo $Y_1^1 = 1$ se sugiere continuar con la prueba 4, Y_2^4 como se observa en el árbol de la Figura 4.11. El resultado en este caso es positivo o negativo de acuerdo a la respuesta obtenida en la segunda aplicación. Si la primera aplicación

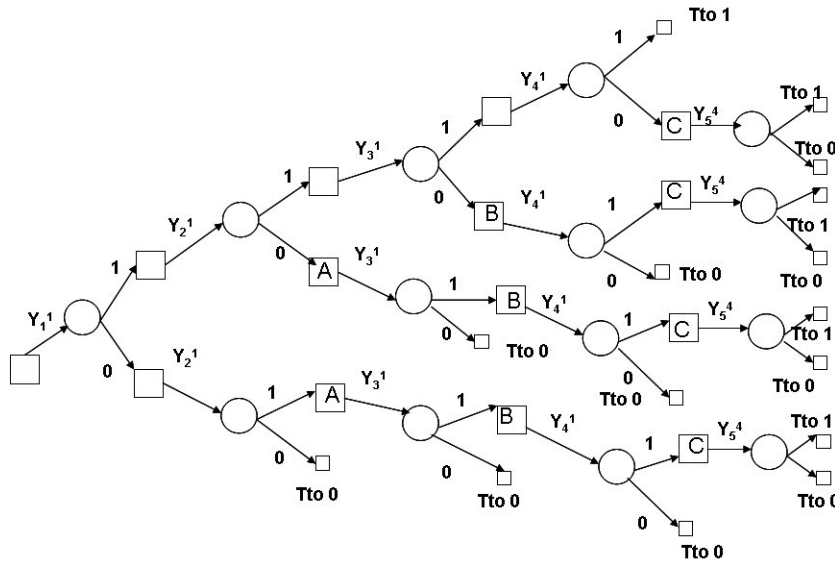


Figura 4.8: Secuencia óptima para las pruebas de monitoreo del ejemplo 2.

de la prueba 1 es negativa se declara negativo al paciente. (Ver archivo [output23.html](#)). Empezar con la prueba 2, que presenta una ganancia cercana 0.00824281 a la óptima 0.00825428 implica utilizar la prueba 4 si $Y_1^2 = 1$ y de su respuesta nuevamente se obtiene la recomendación para el tratamiento. Si $Y_1^2 = 0$ se declara negativo al paciente con una utilidad superior a la obtenida en la secuencia óptima para ésta opción. 0.00900007 por dólar contra 0.000887188 por dólar de la acción óptima. (Ver [output24.html](#))

4.4. Conjuntos de funciones de utilidad

Ahora se tiene un conjunto de funciones de utilidad y el mismo esquema de un problema secuencial con dos acciones: *Seguir* aplicando pruebas y *Detener* el ensayo para recomendar un tratamiento. De nuevo se debe elegir entre las acciones que reporten mejores utilidades. Aplicar el algoritmo *Backward Induction* permite calcular las utilidades de las pruebas consideradas para optar por la máxima utilidad de seguir con el ensayo y a la vez considerando la información obtenida hasta el momento, maximizar la utilidad esperada de aplicar uno de los tratamientos.

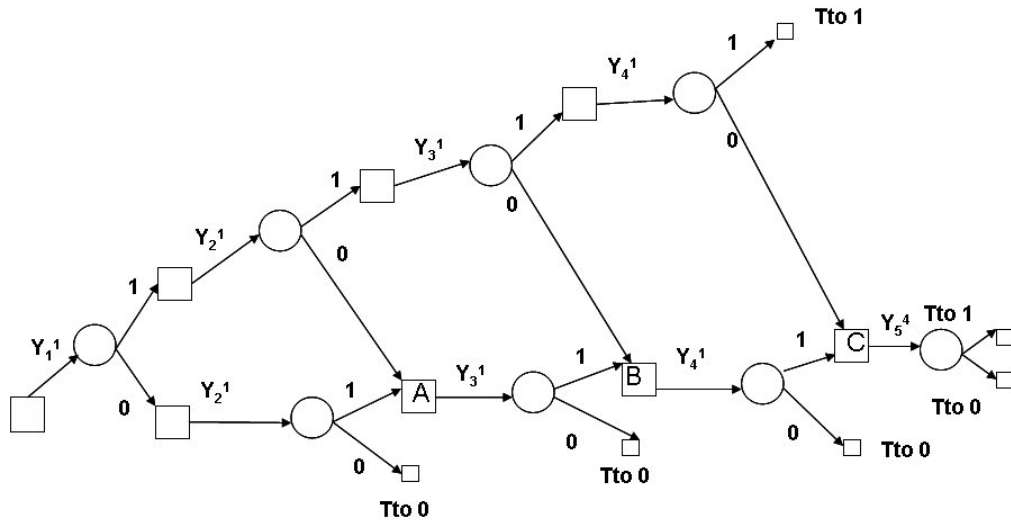


Figura 4.9: Secuencia óptima reducida para las pruebas de monitoreo del ejemplo 2.

La información referente a los conjuntos de utilidades de los tratamientos y costos de las pruebas, aparece ahora en la Tabla 4.4, donde:

$$\begin{aligned}
 q_{00}^{\alpha} &= \alpha q_M(0, 0) + (1 - \alpha) q_m(0, 0), \\
 q_{01}^{\alpha} &= \alpha q_M(0, 1) + (1 - \alpha) q_m(0, 1), \\
 q_{10}^{\alpha} &= \alpha q_M(1, 0) + (1 - \alpha) q_m(1, 0), \quad y \\
 q_{11}^{\alpha} &= \alpha q_M(1, 1) + (1 - \alpha) q_m(1, 1).
 \end{aligned}$$

	$X = 0$	$X = 1$
Tto 0	$[q_m(0, 0), q_M(0, 0)]$	$[q_m(0, 1), q_M(0, 1)]$
Tto 1	$[q_m(1, 0), q_M(1, 0)]$	$[q_m(1, 1), q_M(1, 1)]$

Tabla 4.4: Conjuntos de funciones de utilidad

En la Sección 2.3 se discutió la posibilidad de utilizar una solución gráfica al problema de recomendar un tratamiento cuando la utilidad se

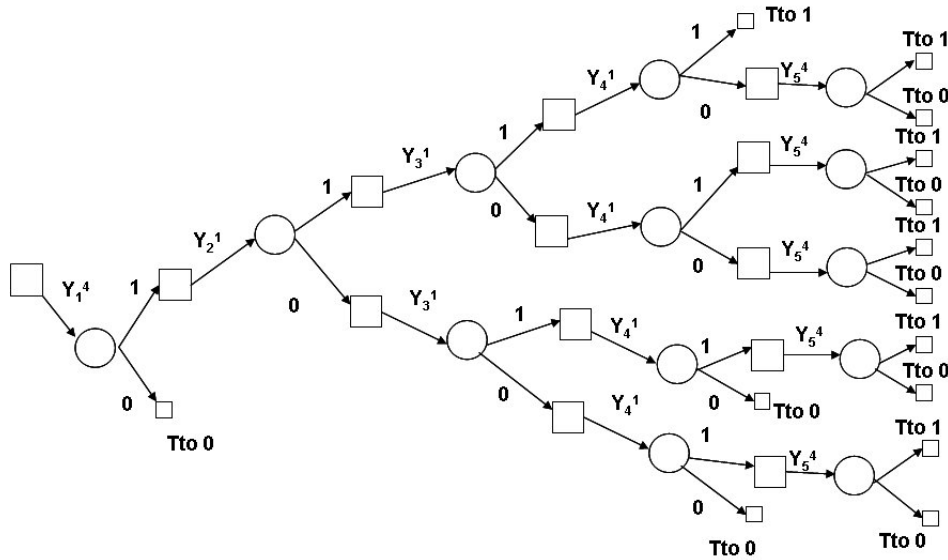


Figura 4.10: Solución al ejemplo 2 de pruebas de monitoreo empezando con la prueba 4.

define como un conjunto de utilidades. Se mencionó además que éste no resultaba un procedimiento óptimo y que podía verse sesgado por cuestiones de apreciación.

Ahora, en un problema de decisión secuencial la dificultad aumenta. Supóngase que se quiere implementar el algoritmo de Backward Induction para resolver el árbol determinado por la situación del ejemplo 2.3. y con la función de utilidad dada por la Tabla 4.4. Para empezar, sea p la probabilidad de tener la condición E en la última etapa del árbol dada la información obtenida en las pruebas realizadas. La ecuación (4.15) muestra las expresiones para obtener la utilidad esperada de cada nodo final:

$$\begin{aligned}
 & \max\{pq_{01}^\alpha + (1-p)q_{00}^\alpha, pq_{11}^\alpha + (1-p)q_{10}^\alpha\} = \quad (4.15) \\
 & \max\{p[\alpha q_M(0,1) + (1-\alpha)q_m(0,1)] + (1-p)[\alpha q_M(0,0) + (1-\alpha)q_m(0,0)], \\
 & \quad p[\alpha q_M(1,1) + (1-\alpha)q_m(1,1)] + (1-p)[\alpha q_M(1,0) + (1-\alpha)q_m(1,0)]\}.
 \end{aligned}$$

Intentar aplicar la idea de calcular las utilidades esperadas para valores extremos de α no funciona en este caso, no se puede garantizar que el máximo

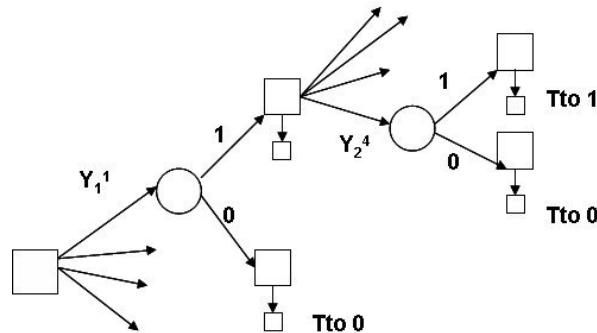


Figura 4.11: Ejemplo de pruebas de monitoreo con un límite de 2 pruebas.

de las utilidades esperadas se pueda escribir como la combinación lineal de los máximos de las utilidades esperadas.

Por ejemplo maximizar las utilidades esperadas en cada tratamiento para $\alpha = 0$ daría la recomendación sobre el mínimo de las utilidades esperadas y para $\alpha = 1$ se tendría la recomendación para el máximo de las utilidades esperadas. Los máximos podrían resultar de tratamientos diferentes y hacer una combinación lineal sobre estas utilidades no correspondería a la realidad, sería un conjunto de utilidades completamente ajeno a las del problema original.

4.5. Recomendaciones

Respecto al caso de las pruebas de diagnóstico de VIH presentadas en la Sección 4.3.1, se observa que entre los precios de las pruebas 1 y 2 hay poca diferencia. Pero la prueba 3 presenta una diferencia bastante mayor, su costo es 10 veces superior al de las pruebas 1 y 2. Por esta razón, se espera tomar una decisión utilizando únicamente las dos primeras pruebas. Las sugerencias de UNAIDS/WHO(2001) van en esta línea, considerando que muchos países no cuentan con los recursos suficientes para aplicar pruebas del tipo 3, Western Blot, que requieren mayor tecnología y personal altamente capacitado. El algoritmo propuesto responde acertadamente a esta necesidad ya que ofrece como secuencia óptima la que implica menores costos y ofrece mejores resultados en el diagnóstico.

En los ejemplos presentados, se observó que la secuencia óptima recomendada como primera aplicación, la prueba 1 y en todos los casos un resultado negativo de ésta primera aplicación sugirió detener la aplicación de pruebas y declarar negativo al paciente, al igual que la estrategia de ONUSIDA pero con la ventaja de emplear sólo las pruebas 1 y 2 para confirmar positivos (ver Figura 4.3), a diferencia de ONUSIDA que sugiere emplear inclusive la prueba 3 aumentando los costos de manera significativa (ver Figura 4.6).

Respecto a las pruebas de monitoreo, se observa una constante referencia al uso de la prueba 1 (ver Figuras 4.8, 4.11 y 4.10), que siendo la más económica constituye un resultado esperanzador. Se debe tener en cuenta sin embargo que hasta el momento la prueba 1 corresponde a una alternativa que no ha sido totalmente estudiada y que su uso dependerá de investigaciones que permitan obtener probabilidades de falsos negativos y positivos más precisas y referidas a diferentes poblaciones.

En el caso de conjuntos de funciones de utilidad es necesario trabajar en el algoritmo para perfeccionar la selección de tratamientos. En un caso real, en que se quiera evaluar varios tratamientos sería de utilidad estudiar la propuesta de tratamientos no dominados que se mencionó en la Sección 2.3. En general, el problema se debe a que en el algoritmo de Backward Induction, la utilidad de una combinación lineal no corresponde a la combinación lineal de las utilidades correspondientes debido a la maximización utilizada.

Es importante buscar la opinión de personas expertas en los procedimientos médicos involucrados en problemas de decisión como los mencionados en el trabajo. Contar con la colaboración de una persona especializada en el tema desde esta perspectiva, aportaría mayor validez al modelo; además de enriquecerlo con sugerencias adicionales o restricciones importantes a considerar en su diseño.

Capítulo 5

Conclusiones

El algoritmo propuesto en este trabajo tiene la ventaja de entregar una respuesta justificada estadísticamente y de proporcionar una respuesta precisa e independiente de la prevalencia de la enfermedad, además de una recomendación en cuanto al tratamiento a aplicar. Las estrategias de ONUSIDA, por su parte, no están respaldadas por un procedimiento matemático y pueden resultar ineficientes para definir el estado de un paciente puesto que puede clasificarlo como indeterminado.

De acuerdo a los datos analizados en el caso de las pruebas diagnóstico, bajo las limitaciones causadas por la falta de información disponible, se observa una constante referencia al uso de las pruebas 1 y 2, las pruebas rápidas, $t = 2$ y pruebas de ELISA, $t = 1$. Estas pruebas difieren poco en sus costos en comparación a la prueba 3, que es 10 veces superior. Por esta razón se espera poder decidir empleando únicamente las dos primeras pruebas. Este objetivo se cumple a cabalidad con el uso del algoritmo ya que las secuencias óptimas que calcula garantizan el diagnóstico del VIH usando sólo las pruebas 1 y 2.

Respecto a las pruebas de monitoreo, el mayor logro se refiere al hecho de presentar una herramienta de decisión. Se mencionó en el Capítulo 4 que no existe una estrategia definida para hacer el seguimiento del SIDA, pero sí el interés por adecuar las recomendaciones y lineamientos empleados por países de altos recursos en salud a otros con menos recursos.

Contar con los valores reales de utilidades propuestas por pacientes y

médicos de VIH y SIDA sería de gran ayuda para realizar estudios más completos y veraces, que ayudarían a responder a las necesidades de un sector específico. En el trabajo se mencionó la existencia de cuestionarios diseñados para calcular utilidades en términos cualitativos de personas que sufren esta enfermedad. No se cuenta con la información sobre su aplicación concreta pero de existir este tipo de base de datos, sería de interés aplicarla y analizar los resultados obtenidos.

La toma de decisiones sobre una estrategia de pruebas debe enmarcarse además en la normatividad del país que la aplique. Considerar sus aspectos éticos y sociales, además de la población que se verá afectada por la política de salud tomada. Sugerencias para la prevención y control de la infección por VIH en México se encuentran en las Normas Oficiales (CONASIDA, 1995).

El algoritmo propuesto debe considerarse como una herramienta adicional a la hora de tomar decisiones y establecer políticas de atención. No pretende ser exclusiva pero sí útil y completa. Lo es, en la medida en que reúne la información disponible respecto a las condiciones de las pruebas, sus costos y funciones de utilidades, en términos de años de vida de calidad, de las personas involucradas. El incluir los aspectos epidemiológicos y económicos del problema lo hacen más específico que las estrategias usuales de aplicación de pruebas empleadas en la toma de decisiones.

Podría resultar en este sentido una herramienta valiosa para los tomadores de decisiones y evaluadores de programas de salud para la atención del VIH/SIDA. Es importante recalcar que el algoritmo propuesto, puede ajustarse a las necesidades económicas específicas de un sector. Siempre propone la estrategia óptima en términos de la mayor ganancia posible (en términos de años de vida por cada dólar invertido) pero a la vez arroja las salidas posibles de optar por otras alternativas disponibles.

Una de las recomendaciones iniciales desde la perspectiva médica, realizada por investigadores del Instituto Nacional de Salud Pública, es considerar el tiempo de aplicación entre pruebas. Las pruebas de monitoreo deben hacerse con una frecuencia mínima de 2 al año para verificar el estado del paciente y tomar decisiones oportunas. En la aplicación del modelo se pueden representar escenarios que varíen los tiempos de aplicación entre pruebas. El efecto de estas variaciones se incluiría en los costos y utilidades propuestos,

para que al realizar una comparación de costo-efectividad entre escenarios permitan establecer una frecuencia óptima de aplicación de las pruebas. Este punto queda abierto para una investigación posterior.

Otra sugerencia se refiere a la necesidad de buscar umbrales óptimos de clasificación para todas las pruebas de monitoreo. Es decir, establecer valores para los cuales las pruebas más económicas puedan indicar el estado de un paciente y declararlo con mayor certeza. De ahí sería posible encontrar equivalencias entre las pruebas. Por ejemplo, se podría encontrar valores determinados para los que un resultado en una prueba de linfocitos totales indicara lo mismo que cierto valor de una prueba de CD4 y con la misma efectividad.

Con la idea de apoyar el uso de pruebas más económicas, se puede pensar en incluir rangos para las pruebas de monitoreo que determinen el nivel de avance de la enfermedad. Esto podría lograrse refinando las salidas de cada prueba empleada considerando subniveles del estado de enfermedad. Como en la solución del problema se tomaron dos salidas posibles: Tener la condición y no tenerla, y estas salidas están determinadas por valores específicos de los resultados de las pruebas, por ejemplo en el caso de las pruebas de monitoreo un resultado inferior a 250 células por cc indica la necesidad de iniciar un tratamiento antiretroviral, se puede pensar en definir más salidas que consideren los niveles determinantes del avance de la enfermedad dados por las pruebas de CD4 y Carga Viral. De ahí sería posible comparar las utilidades de todas las secuencias y definir equivalencias entre pruebas.

Los dos últimos párrafos dejan abiertas posibilidades de investigaciones posteriores que pueden resultar de interés tanto a profesionales de la estadística como de la medicina y la salud pública. Además debe recordarse que la posibilidad de contemplar conjuntos de funciones de utilidad y construir un algoritmo con propiedades similares al de Backward Induction conforma un campo de investigación novedoso y con múltiples aplicaciones.

Finalmente, es de mencionar que el presente trabajo es un claro ejemplo de la contribución de la estadística a la solución de problemas reales. Se sabe que existe la posibilidad de dar mayor validez y significancia al modelo propuesto para la solución de los problemas de VIH y SIDA y que sus resultados podrían extenderse a problemas similares en otros contextos de salud pública o fuera

de éste, pero para hacerlo es necesario un esfuerzo conjunto con investigadores del área y tomadores de decisiones dispuestos a contemplar posibilidades diferentes de acción.

Apéndice A

Anexos

A.1. Ejemplo de aplicación del algoritmo en la Sección 4.2.1.

- Dec, dat=((0, 0) (0, 0)), U=0.192511, cont. with t= 1
 - Rand, t=2, prob=(0.9386, 0.0614,), U=0.161163
 - Dec, dat=((0, 0) (1, 1)), U=0.11886, cont. with t= 1
 - Rand, t=2, prob=(0.1844, 0.8156,), U=0.104629
 - Last, dat=((0, 0) (2, 2)), U=0.0969944, Decision= 1
 - Last, dat=((0, 0) (2, 1)), U=0.138398, Decision= 0
 - Rand, t=1, prob=(0.2130, 0.7870,), U=0.11886
 - Last, dat=((1, 1) (1, 1)), U=0.112446, Decision= 1
 - Last, dat=((1, 0) (1, 1)), U=0.142552, Decision= 0
 - Dec, dat=((0, 0) (1, 0)), U=0.163927, stop, Decision=0
 - Rand, t=1, prob=(0.9329, 0.0671,), U=0.192511
 - Dec, dat=((1, 1) (0, 0)), U=0.147352, cont. with t= 1
 - Rand, t=2, prob=(0.2805, 0.7195,), U=0.126053
 - Last, dat=((1, 1) (1, 1)), U=0.112446, Decision= 1
 - Last, dat=((1, 1) (1, 0)), U=0.160961, Decision= 0
 - Rand, t=1, prob=(0.3047, 0.6953,), U=0.147352
 - Last, dat=((2, 2) (0, 0)), U=0.133676, Decision= 1
 - Last, dat=((2, 1) (0, 0)), U=0.178556, Decision= 0
 - Dec, dat=((1, 0) (0, 0)), U=0.195759, stop, Decision=0

Figura A.1: Desarrollo completo del árbol de decisión correspondiente al caso $T=2$, $N=2$. Esta es la salida del programa en C++ que corresponde al archivo output10.html. Al correr el programa en el computador, se puede observar el desarrollo completo de un nodo al hacer click sobre él. Ver manual de operación en el Anexo B.

A.2. Salida completa del Ejemplo de aplicación de la Sección 4.2.1.

```

> esI
      Nodo      F(X=1|D)      Util0      Util1      Utilfin Dec
[1,] 1 9.918546e-001      2.065163      6.951123      0.13367553 1
[2,] 2 8.938859e-002      9.284891      1.536332      0.17855560 0
[3,] 3 9.952775e-001      2.027780      6.971665      0.11244621 1
[4,] 4 2.555778e-003      9.979554      1.015335      0.16096054 0
[5,] 5 8.938859e-002      9.284891      1.536332      0.17855560 0
[6,] 6 7.912743e-005      9.999367      1.000475      0.19229552 0
[7,] 7 1.452246e-001      8.828203      1.871248      0.14255166 0
[8,] 8 2.065599e-006      9.999983      1.000012      0.16129006 0
[9,] 9 9.952775e-001      2.027780      6.971665      0.11244621 1
[10,] 10 1.452246e-001      8.828203      1.871248      0.14255166 0
[11,] 11 9.972660e-001      2.021872      6.983596      0.09699429 1
[12,] 12 4.415216e-003      9.964678      1.026491      0.13839821 0
[13,] 13 2.555778e-003      9.979554      1.015335      0.16096054 0
[14,] 14 2.065599e-006      9.999983      1.000012      0.16129006 0
[15,] 15 4.415216e-003      9.964678      1.026491      0.13839821 0
[16,] 16 5.391784e-008      10.000000      1.000000      0.13888888 0
> esII
      Nodo      F(Y_2=1|D)      Utilfin
[1,] 17 0.70392740686      0.1469628
[2,] 18 0.72472578241      0.1258010
[3,] 19 0.00200000000      0.1922680
[4,] 20 0.00205906814      0.1612515
[5,] 21 0.79952567237      0.1184816
[6,] 22 0.82313691932      0.1042172
[7,] 23 0.00005230917      0.1612900
[8,] 24 0.00005285294      0.1388889
> esIII
      Nodo      F(X=1|D)      Util0      Util1      Utilfin Dec
[1,] 25 0.71684053651      0.08363286      0.10394202      0.10394202 1
[2,] 26 0.00203665988      0.19575896      0.01984745      0.19575896 0
[3,] 27 0.81418092910      0.05715660      0.09647681      0.09647681 1
[4,] 28 0.00005226799      0.16392744      0.01623868      0.16392744 0
> es
      Nodo      UtilP1      UtilP2      UtilStop      Utilfin Prueb
[1,] 29 0.1469628      0.1258010      0.10394202      0.1469628 1
[2,] 30 0.1922680      0.1612515      0.19575896      0.1957590 5
[3,] 31 0.1184816      0.1042172      0.09647681      0.1184816 1
[4,] 32 0.1612900      0.1388889      0.16392744      0.1639274 5
> esIV
      Nodo      F(Y_1=1)      Utilfin
[1,] 33 0.06710      0.1924847
[2,] 34 0.06135      0.1611393
> esV
      Nodo      F(X=1)      Util0      Util1      Utilfin Dec
[1,] 35 0.08      0.192      0.026      0.192 0
> u36
      Nodo      Utilfin      Prueb
[1,] 36 0.1924847 1

```

Figura A.2: Salida en SPlus del ejemplo de aplicación del algoritmo en la Sección 4.2.1.

Apéndice B

Anexos

B.1. Uso del programa

El programa elaborado en C++ cuenta con dos archivos básicos, input, para ingresar la información del problema y output que recoge las salidas del árbol correspondiente al problema resuelto por Backward Induction como se explicó en el Capítulo 4. En éste anexo se muestra la forma en que debe ingresarse la información al programa así como la interpretación que debe hacerse de su salida. Inicialmente se presenta una descripción general de los datos que se requieren y cómo podrían organizarse, en seguida se muestra el desarrollo del problema descrito en la Sección 4.2.1

La Tabla B.1 recoge la información de costos y utilidades respectivas para el caso de una sola función de utilidad posible para el caso de T pruebas.

	Y^1		Y^2		...	Y^T	
	X=1	X=0	X=1	X=0	...	X=1	X=0
Tto 0	(c_1, q_{11})	(c_1, q_{10})	(c_2, q_{11})	(c_2, q_{10})	...	(c_T, q_{11})	(c_T, q_{10})
Tto 1	(c_1, q_{21})	(c_1, q_{20})	(c_2, q_{21})	(c_2, q_{20})	...	(c_T, q_{21})	(c_T, q_{20})

Tabla B.1: Costos y utilidades para T pruebas

La idea es utilizar la información anterior además de la información a priori e ingresarla al programa en el archivo input que aparece como se muestra en la Figura B.1. Para el ejemplo mencionado, descrito por el árbol

T: Número de pruebas a utilizar. Son 3 para las pruebas diagnósticas y 4 para las de monitoreo.

N: Número máximo de aplicaciones de las pruebas a usar.

p: $P(X=1)$ Probabilidad a priori de tener la enfermedad.

ps: $P(Y^1 = 1|X = 0) = P_{11}, P(Y^1 = 0|X = 1) = P_{01}, P(Y^2 = 1|X = 0) = P_{12}, P(Y^2 = 0|X = 1) = P_{02}, \dots, P(Y^T = 1|X = 0) = P_{1T}, P(Y^T = 0|X = 1) = P_{0T}$

Probabilidades de falsos positivos y falsos negativos de cada una de las pruebas.

C: Costo del tratamiento, por ejemplo el máximo de los dos.

$c : c_1, c_2, \dots, c_T$: costos de las pruebas.

us : $q_{00}, q_{01}, q_{10}, q_{11}$: Utilidades de los dos tratamientos en caso positivo y caso negativo

Figura B.1: Ingreso de información al archivo input

de la figura 4.1, la información se ingresa como aparece en la Figura B.2.

La ejecución del programa se efectúa en una terminal de Linux con la orden *recbi input* y la salida aparece en un archivo de html como el de la figura B.3. En ella se puede seguir la secuencia de pruebas realizadas y sus salidas respectivas. La salida muestra una serie de subniveles que representan la dependencia del resultado anterior, así la cuarta y quinta líneas de la salida hacen parte de la tercera, es decir de optar por iniciar con la prueba 1.

La primera línea indica la primera prueba a aplicar. Aquí es posible que se recomiende no aplicar ninguna prueba, en ese caso se indicaría parar y el tratamiento a seguir. Aparece el conjunto $((0, 0), (0, 0))$ en el que la primera

T: 2
N: 2
p: 0.05
ps:0.02 0.038 0.012 0.001
C: 50
c: 1 11
us:10 2 1 7

Figura B.2: Información del ejemplo 4.2.1

pareja corresponde a la primera prueba, su primera coordenada a las pruebas que se han realizado de éste tipo, y su segunda coordenada a las pruebas que resultan positivas. En forma similar la segunda pareja corresponde a las aplicaciones de la segunda prueba y su total de respuestas positivas. En seguida aparece una utilidad correspondiente a la acción óptima de iniciar con la prueba 1.

Las siguientes líneas corresponden al cálculo de las probabilidades predictivas negativa y positiva de cada una de las pruebas bajo la información obtenida con sus respectivas utilidades. Por ejemplo para la prueba 1, que es la recomendación óptima, se tiene una probabilidad de 0.9329 de obtener un resultado negativo y de 0.0671 de obtener un positivo. Junto a las probabilidades aparece nuevamente la utilidad de aplicar la prueba 1 inicialmente. En seguida en un nivel inferior, aparecen las posibles respuestas de la primera aplicación de la prueba 1. Si $Y_1^1 = 1$ representado por las coordenadas $((1, 1), (0, 0))$ se tiene una utilidad de 0.147352QALYs por dólar y se recomienda aplicar la prueba 1, nuevamente. En un nivel inferior aparecen las posibles salidas de aplicar por segunda vez la prueba 1. Si $Y_2^1 = 1$, se tiene el conjunto $((2, 2), (0, 0))$ junto a la palabra *last* que indica es la última salida (porque el límite de pruebas era 2, $N=2$). Se tiene en este caso una utilidad de 0.133676QALYs por dólar y la recomendación de aplicar el tratamiento 1. La siguiente línea corresponde al caso en que la segunda aplicación de la prueba 1 fue negativa.

La última línea corresponde al caso en que la primera aplicación de la prueba 1 resulta negativa, $((1, 0), (0, 0))$ en cuyo caso se obtiene una utilidad de 0.195759QALYs por dólar y se recomienda aplicar el tratamiento 0.

- Dec, dat=((0, 0) (0, 0)), U=0.192511, cont. with t= 1
 - Rand, t=2, prob=(0.9386, 0.0614,), U=0.161163
 - Rand, t=1, prob=(0.9329, 0.0671,), U=0.192511
 - Dec, dat=((1, 1) (0, 0)), U=0.147352, cont. with t= 1
 - Rand, t=2, prob=(0.2805, 0.7195,), U=0.126053
 - Rand, t=1, prob=(0.3047, 0.6953,), U=0.147352
 - Last, dat=((2, 2) (0, 0)), U=0.133676, Decision= 1
 - Last, dat=((2, 1) (0, 0)), U=0.178556, Decision= 0
 - Dec, dat=((1, 0) (0, 0)), U=0.195759, stop, Decision=0

Figura B.3: Secuencia Óptima

Bibliografía

- [1] Abellan, J.M. and Pinto J.L. (2000). Quality Adjusted Life Years as expected utilities. *Spanish Economic Review*.**2(1)**,49-63.
- [2] Bautista, S.,Dmytraczenko, T., Kombre, G. y Bertozzi S.(2003). Análisis de los costos de atención de VIH/SIDA en México. *Reporte técnico N. 0.20, Bethesda MD:Socios para la reforma del sector salud*
- [3] Bernardo, J.M. and Smith,A. (1994). *Bayesian Theory*. John Wiley & Sons.
- [4] CONASIDA. (1995). Norma Oficial Mexicana para la prevención y control de la infección por VIH. NOM-010-SSA2-1993. *Diario Oficial de la Federación*. <http://www.salud.gob.mx/conasida>
- [5] Christen, J.A. (2004). Notas del Curso de Estadística Bayesiana en CIMAT. Semestre agosto-diciembre 2004.
- [6] Christen, J.A., Müller, P., Wathen, K. and Wolf, J. (2004). A Bayesian Randomized Clinical Trial: A Decision Theoretic Sequential Design. *Canadian Journal of Statistics*, **32(4)**,387-402.
- [7] DELIVER. (2002). Logistics for health. <http://deliver.jsj.com/2002/HIVAIDS/tests/faq/index.cfm>
- [8] Díaz Torres, H. y Lubián, A. (1998). Definición de caso y clasificación de la infección por VIH y SIDA. *Revista Cubana de Medicina*, **37(3)**,157-165.
- [9] Freedberg, K.A., Scharfstein, J., Seage, G., Losina, E., Weinstein, M., Craven, D., and Paltiel, D.(1998). The cost effectiveness of preventing AIDS-related opportunistic infections. *Journal of American Medical Association*. **279(2)**, 130-136.

- [10] Freedberg, K.A., Losina, E., Weinstein, M., Paltiel, D., Cohen, C., Seage, G., Craven, D., Zhang, H., Kimmel, A. and Goldie, S. (2001). The cost effectiveness of combination antiretroviral therapy for HIV disease. *The New England Journal of Medicine*, **344(11)**, 824-831.
- [11] Jacobson, M., Liu L., Khayam-Bashi H., Deeks S., Hecht F. and Kahn J. (2003). Absolute or total lymphocyte count as a marker for the CD4 T lymphocyte criterion for initiating antiretroviral therapy. *Research Letters*.
- [12] Lenert, L. and Kaplan, R. (2000). Validity and interpretation of preference-based measures of Health-related quality of life. *Medical Care*, **38(9),supp. II**,138-150.
- [13] McGregor, M. (2003). Cost-Utility analysis: Use QALYs only with great caution. *Health Economics*. **18**
- [14] Miró, J., Sued, O., Plana, M., Pumarola, T. y Gallart, T. (2004). Avances en el diagnóstico y tratamiento de la infección aguda por el VIH-1. *Enfermedades Infecciosas y Microbiología Clínica*. **22(10)**,643-659.
- [15] Ortún, V., Meneu, R. y Peiró, S. (2004). El impacto de los servicios sanitarios sobre la salud. *Working Papers, Research Center on Health and Economics*. Department of Economics and Business, Universitat Pompeu Fabra. <http://www.econ.upf.edu/docs/papers/downloads/752.pdf>
- [16] Parmigianni, G. (2002). *Modeling in Medical Decision Making. A Bayesian Approach*. John Wiley & Sons.
- [17] Rios-Insua, D., Martin, J., Proll, L., French, S. and Salhi, A. (1997). Sensitivity analysis in statistical decision theory: a decision analytic view. *Journal of statistical computing and simulation*, **57**, 197-218.
- [18] UCSF. (2002). Puede el análisis de costo-efectividad ayudar en la prevención del VIH? *Center for AIDS prevention studies*. <http://www.caps.ucsf.edu/costo-efectividad.html>
- [19] UNAIDS/WHO. (1997). Revised recommendations for the selection and use of HIV antibody tests. *Weekly Epidemiological Record*, **72**, 81-88.

- [20] UNAIDS/WHO. (2001). Guidelines for using HIV testing technologies in surveillance: selection, evaluation and implementation. *WHO/CDS/CSR/EDC/2001.16*
- [21] UNAIDS/WHO. (2004). HIV Assays Operational Characteristics. *Report 14/Simple/Rapid Test*.
- [22] Velarde-Jurado, E. y Avila-Figueroa, C. (2002). Evaluación de la calidad de vida. *Salud pública de México*.**44(4)**, 349-361.