Reducción de las probabilidades condicionadas en la estimación del esfuerzo software mediante Redes Bayesianas

Joseba Esteban López, José Javier Dolado
Departamento de Lenguajes y Sistemas
Universidad del País Vasco U.P.V./E.H.U.
joseba.esteban@gmail.com, dolado@si.ehu.es

Resumen

La aplicación de las Redes Bayesianas en el campo de la estimación a partir del conocimiento experto se ve obstaculizada por la cantidad de parámetros requeridos para su funcionamiento. El parámetro que más estimaciones requiere por parte del experto son las tablas de probabilidades condicionadas. En este artículo se estudia el método propuesto por Kwai-Sang Chin et al. para la reducción de las probabilidades condicionadas. En base a esta propuesta, se aporta una mejora al mismo mediante el empleo de la distribución de probabilidad Normal. También se presenta una aplicación de esta propuesta al área de la estimación del esfuerzo de proyectos software.

1. Introducción

La estimación del esfuerzo de desarrollo de un proyecto software representa cada vez un papel más importante dentro de la creación y mantenimiento de productos y sistemas software. La estimación del esfuerzo de desarrollo de un proyecto software permite evaluar la viabilidad del proyecto así como analizar sus posibles alternativas y gestionar los recursos necesarios para el desarrollo del mismo. Por tanto, disponer de una buena estimación en las etapas iniciales del desarrollo del proyecto software resulta clave para la compañía desarrolladora. Con este fin se han desarrollado diferentes métodos de estimación a lo largo de las últimas décadas [3] [4]. No obstante, los resultados esperados en cuanto a la exactitud de las estimaciones obtenidas están aún por alcanzar [2]. Esto puede deberse a que la estimación del esfuerzo de desarrollo en las etapas iniciales de un proyecto software debe realizarse teniendo en cuenta una cantidad considerable de aspectos que aportan incertidumbre en las estimaciones (los cambios en el ciclo de vida del proyecto software, el alto grado de novedad que suelen tener estos proyectos, cambios en la plantilla, variación de los requisitos por parte del cliente, etc.) las cuales repercuten negativamente en la exactitud de los estudios precedentes. Estos estudios carecen de la gestión de necesaria para dichas incertidumbres y suelen basar sus estimaciones en fórmulas matemáticas más o menos complejas.

En este punto, las redes Bayesianas se presentan como un método eficaz para la gestión de estas incertidumbres, siendo capaz de representar las relaciones entre los elementos de un proceso de razonamiento [5] [7] [4]. Tanto la estructura de la res como los parámetros necesarios en las redes Bayesianas se pueden aprender a partir de datos almacenados sobre estimaciones anteriores. Mendes et al. realizan un estudio sobre la estimación del esfuerzo software mediante redes Bayesianas a partir de datos históricos [8] [6] obteniendo una baja exactitud en sus resultados. Por otro lado, estas redes probabilísticas también permiten obtener sus parámetros a partir del conocimiento experto, permitiendo, además, manejar la incertidumbre propia del las estimaciones realizadas por el experto. Esto implica que la captura de dicho conocimiento debe realizarse de la forma más reducida posible ya que si, en el proceso de captura del conocimiento, el experto debe estimar una gran cantidad de parámetros, puede derivar en inconsistencias y/o prejuicios que provocarán errores en las estimaciones resultantes y, por tanto, disminuirán la exactitud de las mismas. Los parámetros necesarios en una red Bayesiana tienen forma de probabilidades y son numerosos. La mayor parte de ellos se concentra en la especificación de las tablas de probabilidad condicionadas. Estas tablas se dan cuando un nodo de la red se ve influenciado por otro (representado mediante un arco) y representan la probabilidad de que se den los posibles estados del nodo hijo (nodo influenciado) dados los posibles estados del nodo padre (nodo que ejerce la influencia). Estas tablas de probabilidad muestran la probabilidad de que se de cada uno de los posibles estados del nodo hijo para cada una delas combinaciones entre los estados de los nodos padre.

La cantidad de probabilidades a estimar por el experto para obtener estas tablas de probabilidad condicionadas depende del número de estados del nodo hijo, el número de nodos padre y la cantidad de posibles estados de estos nodos. De esta forma, dado un nodo hijo con cinco posibles estados y 10 padres, cada uno con tres posibles estados, el experto debería proporcionar 295.245 estimaciones $((n^P)h$, siendo n el número de estados de los nodos padre, P el número de padres y h el número de estados del nodo hijo). Esto hace necesaria la reducción de estimaciones en la generación de estas tablas de probabilidad [7]. Chin et al. proponen un método de generación de las tablas de probabilidad condicionadas en [1].

En el segundo apartado de este artículo se presenta la propuesta de Chin . En la siguiente sección se propone una mejora a la propuesta de Chin. A continuación se presenta la aplicación de esta propuesta para la estimación del esfuerzo de desarrollo de proyectos software. Finalmente se presentan las conclusiones de este estudio.

2. Propuesta de Chin: Reducción de las probabilidades condicionadas mediante AHP

En el artículo [1] los autores presentan un método que permite generar las las probabilidades condicionadas de las redes Bayesianas reduciendo el número de estimaciones a proporcionar por el experto. En la fórmula (1) se presenta la fórmula en la que se basa el método propuesto por Chin, presentada por Kim y Pearl en [9]. Esta formula establece que la probabilidad condicionada de un nodo dados sus padres $(P(A|X_1, X_2, ..., X_n))$, puede ser aproximada por el producto de la probabilidad del nodo hijo dado cada unos de sus padres por separado $(\alpha P(A|X_1) * P(A|X_2) * ... * P(A|X_n))$, siendo α un factor de normalización.

$$P(A|X_1, X_2, ..., X_n) = \alpha \\ P(A|X_1) * P(A|X_2) * ... * P(A|X_n) \tag{1}$$

2.1. Generación de las probabilidades condicionadas para un nodo con un único padre

Para la estimación de cada una de las probabilidades condicionadas de un nodo con un único padre, $P(A|X)$, los autores presentan una propuesta. Esta propuesta permite establecer las probabilidades un nodo condicionado por un único padre utilizando la tabla de comparaciones del método de ayuda a la toma de decisiones AHP ([10], [11]).

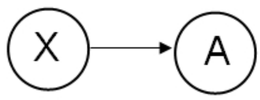

Figura 1: Red Bayesiana con un nodo hijo y un nodo padre

Supongamos la red Bayesiana de la figura

1. El nodo A tiene un único padre, X, con n y m estados respectivamente. Estos estados se pueden representar por $a_1, a_2, ..., a_n$ y $x_1, x_2, ..., x_m$. El objetivo es estimar la probabilidad de cada estado de A para cada uno de los estados del nodo X, por ejemplo $P(A = a_i|X = x_j)$ con $i = 1..n$ y $j = 1..m$. Con el fín de establecer cada una de estas probabilidades, se utiliza la matriz de comparaciones del método AHP. Cada una de las posiciones de la matriz, a_ij con $i = 1..n$ y $j = 1..m$, son estimadas por el experto contestando a preguntas del tipo:

> "Suponiendo que el nodo X toma el valor x_h con $h = (1..m)$, comparando los estados del nodo A, a_i y a_j, ¿cuál es más probable que ocurra y cuánto más?"

En la tabla 1 se muestra la tabla de comparaciones correspondiente cuando el nodo X toma el estado x_h con $h = 1..m$). Una vez rellenada la tabla mediante las comparaciones realizadas por el experto, se puede calcular la prioridad o peso de cada uno de los estados dado $X = x_h$, $\omega_h = (\omega_{1h}, \omega_{2h}, .., \omega_{nh})^T$, mediante el cálculo del máximo *eigenvector* [10].

$X = x_h$	a_1	a_2	...	a_n	ω_h
a_1	a_{11}	a_{21}	...	a_{n1}	ω_{1h}
a_2	a_{12}	a_{22}	...	a_{n2}	ω_{2h}
...	...	...	...	...	...
a_n	a_{1n}	a_{2n}	...	a_{nn}	ω_{nh}

Cuadro 1: Matriz de comparaciones y pesos dado un estado del nodo padre: $X = x_h$

Debido a que la suma de los elementos del vector ω_h es 1 y que que cada ω_{ih} representa la probabilidad de ocurrencia relativa de cada posible estado de A, a_i, dado que el nodo X toma el estado x_h, se puede interpretar ω_{ih} como la probabilidad de que se de a_i dado que el nodo X toma el estado x_h: $P(A = a_i|X = a_h) = \omega_{ih}$. De este modo, con el objetivo de completar la tabla de probabilidades condicionadas entre los nodos X y A (ver tabla 1), el experto debe estimar una matriz de comparaciones por cada estado que puede tomar nodo padre, X.

3. Independencia del número de estados del nodo hijo: Distribución Normal

Empleando el método de generación de las probabilidades condicionadas propuesto por Chin et al., el número de estimaciones que debe realizar el experto depende de tres valores:

- La cantidad de nodos padre
- El número de posibles estados de cada padre
- El número de estados del nodo hijo

En este apartado se propone un método basado en las características de la distribución Normal para la generación de las probabilidades condicionadas independiente del número de estados del nodo hijo. Con el fin de entender mejor esta propuesta, este apartado se divide en dos secciones. En la primera de ellas se presentan los aspectos más importantes de las distribuciones Normales y el enfoque que se ha aplicado para establecer las probabilidades de los posibles estados de un nodo. El la segunda parte se expone la propuesta para reducir el número de estimaciones necesarias por el método de Chin.

3.1. Distribución Normal: Cálculo de probabilidades de los estados de un nodo

La distribución Normal es una de las distribuciones de probabilidad para variables aleatorias continuas. La distribución Normal tiene forma acampanada y simétrica respecto a uno de sus parámetros: la media, μ. La normal tiene dos parámetros: la media, μ, y la desviación típica, σ. La media representa el valor más probable y la desviación típica representa la media de distancias que tienen los datos respecto de su media aritmética. El parámetro σ es una medida de dispersión respecto de μ de los datos representados en la distribución normal.

Dado que el objetivo es capturar el juicio experto, se puede interpretar μ como el valor estimado por el experto, mientras que σ como la incertidumbre del juicio aportado por el experto, es decir, una medida de dispersión del juicio efectuado por el experto. De esta forma, la normal permite modelar las estimaciones realizadas por el experto. La media estará representada por un valor estimado por el experto entre 0 y 100, y la desviación típica representará la incertidumbre asociada a los juicios que efectúe.

El cálculo de σ no es tan directo como el de la media. Con el objetivo de establecer su valor, primero es necesario conocer el grado de confianza depositado en las estimaciones realizadas por el experto y, a partir de ahí, calcular σ.

El grado de confianza depositada en las valoraciones realizadas por el experto o la fiabilidad de las mismas, puede modelarse como la calidad o el nivel que tenga el experto a la hora de realizar estimaciones dentro de un área determinada. En el caso de la estimación del esfuerzo software, se pude establecer este índice como la suma de la experiencia del experto dentro la ingeniería del software (en años),ExI, la experiencia dentro de la compañía desarrolladora sobre la que se quiere estimar el esfuerzo (en años), ExC, y el porcentaje de acierto en estimaciones posteriores, P. Estos factores no afectan por igual al resultado objetivo: fiabilidad del experto. Por tanto se le asignará un peso a cada uno de ellos, siendo de 15 % para el primer factor, W_i, y 25 % y 60 % para el segundo y el tercero, W_c y W_p respectivamente. Esta suma ponderada (2) la representaremos por el índice NEst, acrónimo de *Nivel como Estimador*, pero normalizada entre 0 y 100 con el objetivo de obserbar mejor el nivel de fiabilidad que se quiere obserbar (3). Este indicador puede ser calculado de diversas formas, pudiendose adaptar a los factores que más afectan a la fiabilidad del experto dentro del área donde se pretende realizar la estimación.

$$S = (ExI * W_i) + (ExC * W_c) + (P * W_p) \quad (2)$$

$$NEst = \frac{S * 100}{7880} \quad (3)$$

A partir del calculo del grado de fiabilidad depositado en el experto se puede establecer la desviación típica (σ) correspondiente. Con este fin, tendremos presente una de las características de las distribuciones normales: el 99.8 % de los datos representados en una distribución normal se encuentra entre $\mu - 3\sigma$ y $\mu + 3\sigma$, es deir, el 99.8 % de los datos representados en una distribución normal se encuentran en un radio de 3σ respecto de la media, μ, (ver figura 2). Dado que los valores de μ son estimados por experto entre 0 y 100, se puede establecer que $3\sigma = 100$, siendo 100 el máximo diametro establecido en la normal. Cuanto menor sea el diámetro, menor será el valor de la desviación típica y, por tanto, representará una distribución normal con menor incertidumbre. De este modo se puede establecer la desviación típica dependiendo del grado de fiablidad depositado en el experto (4). Así, cuanto mayor sea la fiabilidad del experto (mayor valor de NEst), menor será σ y, por tanto, menor será la incertidumbre representada en la distribución normal, con lo que se otorga mayor confianza en el valor estimado por el experto, μ. Del mismo modo, cuanto menor sea el valor del coeficiente NEst, mayor será el valor de σ y, por tanto, la confianza depositada enla estimación del experto, μ, será menor.

$$\sigma = \frac{100 - NEst}{3} \quad (4)$$

Una vez obtenida σ podemos representar las estimaciones realizadas por el experto mediante la $Normal(\mu, \sigma)$. Esta distribución

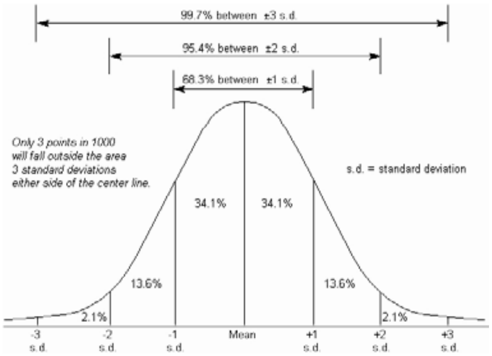

Figura 2: Distribución Normal: 99.8 % de los datos en un radio de 3 σ

es continua. Con el objetivo de poder obtener una probabilidad para cada uno de los estados de puede tomar la variable aleatoria, la distribución normal se discretiza utilizando la función de distribución acumulada (ver figura 3).

$$\Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.$$

Figura 3: Función de Distribución de la Normal acumulada

La distribución normal se puede dividir en tantas partes iguales como sean necesarias. Si se establece que la variable aleatoria (X) puede tomar 3 estados (*Low*, *Average* y *High*), la probabilidad de que la variable aleatoria tome el estado *Low* vendrá dada por la suma de las probabilidades en el intervalo $(0, 33]$, probabilidad de *Average* las del intervalo $(33, 66]$, y la probabilidad de *High* por las probabilidades contenidas en el intervalo $(66, 100]$. Las probabilidades de los estados se X se calculan del siguiente modo:

- $P(X = Low) = \phi(33) - \phi(0)$
- $P(X = Average) = \phi(66) - \phi(33)$
- $P(X = High) = \phi(100) - \phi(66)$

3.2. Generación de probabilidades condicionadas

A partir del planteamiento presentado en el apartado anterior, se puede establecer las probabilidades condicionadas de un nodo hijo dados sus padres, independientemente del número de posibles estados definidos para el nodo hijo. Este proceso se basa en el proceso de generación de las probabilidades condicionadas aportado por Chin et al. [1], expuesto en la sección 2, y en la generación de las probabilidades a partir de la distribución normal.

Como propone Chin et al. en [1], la probabilidad condicionada de un nodo dados sus padres ($P(A|X_1, X_2, ..., X_n)$), se aproxima por el producto de la probabilidad del nodo hijo dado cada unos de sus padres por separado ($\alpha P(A|X_1) * P(A|X_2) * ... * P(A|X_n)$), siendo α un factor de normalización.

En este caso el cálculo de la probabilidad del nodo A condicionado por cada uno de sus padres, $P(A|X_h)$ con $h = 1..n$), se obtiene utilizando la matriz de comparaciones propuesta en el método de ayuda a la toma de decisiones, AHP. A diferencia de la propuesta de Chin, en este caso el experto deberá responder a preguntas del tipo:

> "Comparando los estados de X_h, x_i y x_j con $h = 1..n$, $i = (1..m), j = (1..m)$ y $i \neq j$, ¿cuál de ellos supone un mayor valor de A y cuánto más?"

Contestando a esta pregunta el experto puede rellanar la matriz de comparaciones (2), a partir de la cual, al igual que con la propuesta de Chin, se puede obtener el peso o prioridad de cada estado de X_h: $\omega_{AX_h} = (\omega_{AX_h1}, \omega_{AX_h2}, .., \omega_{AX_hm})^T$, siendo m el número de posibles estados del nodo X_h. Cada ω_{AX_hi} con $i = 1..m$, representa el valor numérico más probable que tomará el nodo A dado cada uno de los estados del nodo padre, X_hi.

Para poder establecer una probabilidad a cada estado de A se aplica los expuesto en

X_h	x_1	x_2	...	x_m	ω_{AX_h}
x_1	x_{11}	x_{21}	...	x_{m1}	ω_{AX_h1}
x_2	x_{12}	x_{22}	...	x_{m2}	ω_{AX_h2}
...	...	...	...	...	...
x_m	x_{1m}	x_{2m}	...	x_{mm}	ω_{AX_hm}

Cuadro 2: Pesos y matriz de comparaciones del valor más probable que toma en nodo hijo A dado cada estado del nodo padre: X_{hi}

el apartado 3.1. En este caso se define una normal por cada ω, coincidiendo el valor de ω con el de la media, $\mu_i = \omega_i$ con $h = 1..m$. La desviación típica es propia del experto que realiza las estimaciones y se calcula mediante la formula (4) expuesta en el apartado 3.1.

Una vez definidas las normales para cada ω_i ($N(\omega_1, \sigma)$, $N(\omega_2, \sigma)$, ... , $N(\omega_m, \sigma)$), bastará con discretizar cada normal en tantos intervalos como estados tenga el nodo hijo A. El proceso de discretización se realiza según lo expuesto en el apartado 3.1. Una vez finalizado este proceso obtendremos las probabilidades condicionadas del nodo A dados cada uno de sus padres $X_1, X_2, ..., X_m$ por separado.A partir de este punto, se puede continuar con el proceso para la generación de la tabla de probabilidades condicionadas expuesto por Chin et al. en [1].

Este método de generación de las probabilidades condicionadas sólo depende del número de padres y del número de estados de estos, siendo independiente del número de estados del nodo hijo. Esto se debe a que la normal estimada para cada estado de los nodos padre, se puede discretizar en tantos intervalos como nodos tenga el nodo hijo.

4. Aplicación en la estimación del esfuerzo del software

Lo expuesto en este artículo se puede aplicar a la estimación del esfuerzo de desarrollo del software. Con el objetivo de ilustrar mejor este concepto, a continuación se presenta un ejemplo de como generar las probabilidades condicionadas en una red Bayesiana que permita estimar del esfuerzo de desarrollo de un proyecto software.

En este caso, el esfuerzo de desarrollo está representado por el nodo *Ef*, con cinco posibles estados: *Very Low* (*VL*), *Low* (*L*), *Average* (*A*), *High* (*H*) y *Very High* (*VH*). El esfuerzo de desarrollo del software está influenciado por múltiples factores. En este ejemplo y con el propósito de no extender demasiado este estudio, supondremos que el esfuerzo se ve influenciado por tres factores (la experiencia de la plantilla (Ex), la complejidad del proyecto (C) y la magnitud del proyecto (M)) cada uno de ellos con tres posibles estados: *Low* (*L*), *Average* (*A*) y *High* (*H*). En la figura 4 se muestra la configuración de la red Bayesiana, donde *Ef*, el nodo hijo, esta condicionado por múltiples padres, los nodos *Ex*, *C* y *M*.

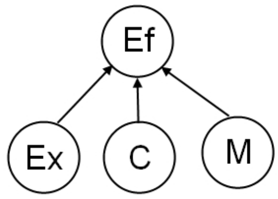

Figura 4: Red Bayesiana para la estimación del esfuerzo Software, *Ef*, a partir de los factores *Ex*, *C* y *M*

Uno de los parámetros necesarios para poder estimar el esfuerzo de desarrollo mediante esta red Bayesiana, es la tabla de probabilidades condicionas entre el nodo hijo, *Ef*, y sus nodos padres, *Ex*, *C* y *M*. Esta tabla establece una probabilidad para cada estado del nodo hijo por cada una de las combinaciones entre los estados de los nodos padre. Dado que, en este caso, el esfuerzo está influenciado por tres nodos (*Ex*, *C* y *M*) cada uno de ellos con tres estados, el número de combinaciones de sus estos es 27 (3^3). Para cada una de estas combinaciones se ha de estimar una probabilidad por cada uno de los cinco estados del nodo hijo, *Ef*. Por tanto el experto debería realizar 135 estimaciones para rellenar la tabla de probabilidades condicionadas. A continuación, se aplica el método para generar la tabla de probabilidades condicionadas reduciendo el número de estimaciones a realizar por el experto.

Lo primero que debe calcular es la fiabilidad que se otorga al experto en las estimaciones que realizará a lo largo de este proceso. Para ilustrar esta idea, supongamos un experto ficticio con los siguientes datos y el peso correspondiente a estos datos:

- Años de experiencia en la ingeniería del software: 5
- Años de experiencia en la compañía desarrolladora del software a estimar: 3
- Porcentaje de acierto en anteriores estimaciones (Pred(25)): 52.67 %

A partir de los datos del experto se le asigna un grado como estimador (índice NEst) calculado mediante las ecuaciones (2) y (3). En este caso $NEst = 42,02$. En la formula (5) se muestra el cálculo de σ a partir del coeficiente NEst obtenido a partir de los datos del experto. Esta desviación típica (σ) se utilizará para todas las estimaciones realizas por el experto.

$$\sigma = \frac{100 - 42{,}02}{3} = 19{,}32 \qquad (5)$$

A continuación el experto debe rellenar una matriz de comparaciones (AHP) por cada factor (nodo padre). Las tablas 3, 4 y 5 muestran las comparaciones realizadas por el experto contestando a preguntas como:

> "Comparando los estados de la *Experiencia*, L y A, ¿cuál de ellos supone un mayor valor del *Esfuerzo* y cuánto más?"

Una vez rellenadas las tablas, se calcula el peso para cada uno de los estados de cada factor (en el caso anterior, para los estados de la *Experiencia* (Ex)), que viene representado a la derecha de la matriz de comparaciones. Estos pesos junto con la desviación típica

Ex	L	A	H	ω_{EfEx}
L	1	1.4	7	0.5385
A	0.71	1	5	0.3846
H	0.14	0.2	1	0.00769

Cuadro 3: Pesos y matriz de comparaciones del valor más probable que toma en nodo *Ef* dados los estados del factor *Ex*

C	L	A	H	ω_{EfC}
L	1	0.57	0.44	0.2
A	1.75	1	0.78	0.35
H	2.25	1.29	1	0.45

Cuadro 4: Pesos y matriz de comparaciones del valor más probable que toma en nodo *Ef* dados los estados del factor C

calculada anteriormente, definen una distribución de probabilidad del *Esfuerzo* para cada estado de los factores *Ex*, C y M, como se puede ver en la tabla 6.

Dado que el nodo hijo (Ef) tiene cinco posibles estados, cada una de estas normales se discretiza en cinco intervalos asociando cada uno de ellos con un estado del *Esfuerzo*. Por ejemplo para el estado L del factor *Ex*, sería:

- $P(X = VL|Ex = L) = \phi(20) - \phi(0)$
- $P(X = L|Ex = L) = \phi(40) - \phi(20)$
- $P(X = A|Ex = L) = \phi(60) - \phi(40)$
- $P(X = H|Ex = L) = \phi(80) - \phi(60)$
- $P(X = VH|Ex = L) = \phi(100) - \phi(80)$

Realizando este proceso para todos cada uno de los estados de los factores, se obtienen las probabilidades $P(Ef|Ex)$, $P(Ef|C)$ y $P(Ef|M)$ (ver tabla 7). Por último sólo queda aplicar la ecuación $P(Ef|Ex, C, M) =$

M	L	A	H	ω_{EfM}
L	1	0.5	0.36	0.1739
A	2	1	0.73	0.3478
H	2.75	1.38	1	0.4783

Cuadro 5: Pesos y matriz de comparaciones del valor más probable que toma en nodo *Ef* dados los estados del factor M

	Estado	μ	σ
Ex	L	0.5385	19.32
	A	0.3846	19.32
	H	0.00769	19.32
C	L	0.2	19.32
	A	0.35	19.32
	H	0.45	19.32
M	L	0.1739	19.32
	A	0.3478	19.32
	H	0.4783	19.32

Cuadro 6: Parámetros de la distribución Normal para cada uno de los estados de los factores Ex, C y M

$\alpha P(Ef|Ex) * P(Ef|C) * P(Ef|M)$ para obtener la tabla de probabilidades condicionadas.

5. Coste en estimaciones de ambas propuestas

Ambos estudios presentados reducen el número de estimaciones a proporcionar por parte del experto. La cantidad de estas estimaciones depende del número de nodos padre (P), el número de estados de los nodos padre (n) y el número de estados del nodo hijo (m). Así las estimaciones a proporcionar por el experto sin utilizar ninguno de estos métodos es $(n^P) * m$.

	Est	VL	L	A	H	VH
Ex	L	0.05	0.18	0.38	0.30	0.09
	A	0.14	0.36	0.35	0.13	0.02
	H	0.57	0.35	0.07	0.01	0
C	L	0.39	0.42	0.17	0.02	0
	A	0.17	0.39	0.32	0.09	0.03
	H	0.1	0.29	0.39	0.19	0.03
M	L	0.43	0.41	0.14	0.02	0
	A	0.18	0.39	0.32	0.1	0.01
	H	0.06	0.26	0.4	0.23	0.05

Cuadro 7: Probabilidades condicionadas del Esfuerzo dados los factores Ex, C y M: $P(Ef|Ex)$, $P(Ef|C)$ y $P(Ef|M)$

Utilizando el método propuesto por Chin, el número de estimaciones se establece a partir del número de estimaciones necesario para rellenar las matrices de comparaciones , $\frac{n(n-1)}{2}$, el número de nodos padre (P) y el número de estados del nodo hijo (m), dado que que hay que rellenar una matiz por cada uno los estados de cada nodo padre. Por tanto, el número total de estimaciones a proporcionar por el experto utilizando en método propuesto por Chin viene dado por $\frac{n(n-1)}{2} * P * m$.

Utilizando el método propuesto basado en la distribución normal, el número de estimaciones a proporcionar por el experto se calcula a partir del número de estimaciones necesarias para rellenar las matrices de comparaciones, $\frac{n(n-1)}{2}$, y el número de nodos padre (P). El número total de estimaciones necesarias utilizando este método se calcula aplicando $\frac{n(n-1)}{2} * P$.

En las figuras 5, 6, 7 y 8 se muestra el grado de reducción de las estimaciones necesarias para rellenar las tablas condicionadas con 3, 5 y 10 nodos padres para cada uno de los métodos presentados. En la figura 5 el nodo hijo tiene tres estados ($m = 3$) y los nodos padres pueden tomar tres posibles estados ($n = 3$). En la figura 6, $m = 3$ y $n = 5$. El la figura 7, $m = 5$ y $n = 3$. Y en la figura 8, $m = 5$ y $n = 5$.

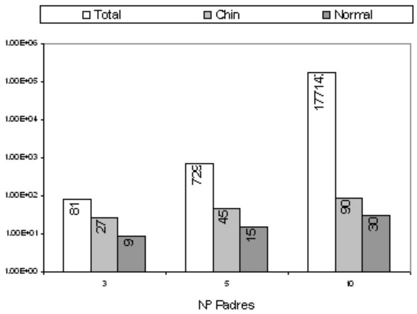

Figura 5: Reducción del número de estimaciones con 3, 5 y 10 nodos padre, $n = 3$ y $m = 3$, aplicando el método de Chin y el de la Normal

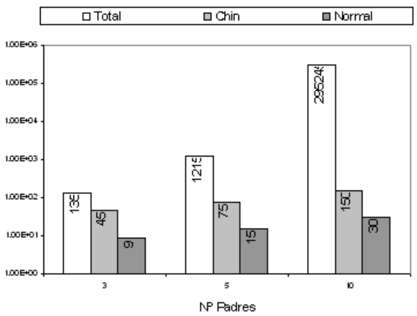

Figura 6: Reducción del número de estimaciones con 3, 5 y 10 nodos padre, $n = 3$ y $m = 5$, aplicando el método de Chin y el de la Normal

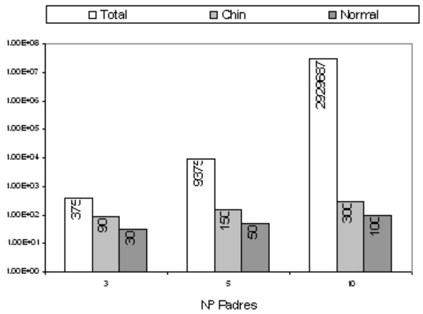

Figura 7: Reducción del número de estimaciones con 3, 5 y 10 nodos padre, $n = 5$ y $m = 3$, aplicando el método de Chin y el de la Normal

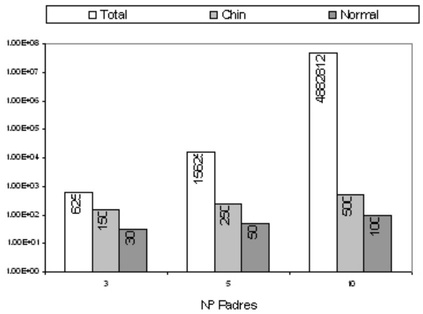

Figura 8: Reducción del número de estimaciones con 3, 5 y 10 nodos padre, $n = 5$ y $m = 5$, aplicando el método de Chin y el de la Normal

6. Conclusiones

En este artículo se ha expuesto uno de los principales problemas que se debe afrontar en la estimación mediante redes Bayesianas a partir del juicio experto: la cantidad de estimaciones necesarias para establecer la tabla de probabilidades condicionadas para nodos con múltiples padres. Con el fin de reducir este problema, se han presentado dos propuestas que permiten disminuir el número de estimaciones que debe aportar el experto.

El método propuesto por Chin et al. en [1] reduce el número de estimaciones pero es dependiente del número de estados del nodo hijo. Esta cantidad de estimaciones es abordable por el experto siempre que el número de nodos padre no sea mayor de cinco. En el caso de la estimación del esfuerzo software, la cantidad de factores (nodos padre) que afectan al esfuerzo es muy amplia, aunque un número factible para su estimación podría ser de 10 factores. En este caso, con el método de Chin y cinco estados tanto para los nodos padre como para el nodo hijo, el número de estimaciones a proporcionar por el experto sería de 48.828.125, sin contar las estimaciones de las probabilidades a priori, lo que hace de este método un sistema poco recomendable para la estimación del esfuerzo software.

El otro método propuesto, basado en el de Chin, permite independizar el número de estimaciones a proporcionar por el experto de la cantidad de estados del nodo hijo. Esto permite reducir más dicho número de estimaciones, haciendo factible la estimación las tablas condicionadas cuando un nodo hijo cuenta con 10 nodos padre (100 estimaciones). Aunque el número de estimaciones requeridas al experto todavía es elevado, este método podría ser utilizado en la estimación del esfuerzo con redes Bayesianas.

Referencias

[1] Kwai-Sang Chin, Da-Wei Tang, Jian-Bo Yang, Shui Yee Wong, Hongwei Wnag, *Assesing new product development project risk by Bayesian network with a systematic probability generation metodology*, Expert Systems with Applications, 2009, 36, 9879 - 9890.

[2] Dolado, J. J. *On the problem of the software cost function Information and Software Technology*, 2001, 43, 61-72

[3] Leung, Z. F. H. *Software Cost Estimation*

[4] Isabel Ramos, Javier Tuya y Javier Dolado. *Técnicas cuantitativas para la gestión en la Ingeniería del Software*, Netbiblo, 2007

[5] Castillo, E.; Gutierrez, J. M. & Hadi, A. S. *Expert Systems and Probabilistic Network Models Springer-Verlag New York*, Inc., 1996

[6] Chulani, S.; Boehm, B. & Steece, B. *Calibrating Software Cost Models Using Bayesian Analysis*, IEEE Transactions on Software Engineering., JulyAugust, 1999, 573-583

[7] Fenton, N. E.; Neil, M. & Caballero, J. G. *Using Ranked Nodes to Model Qualitative Judgments in Bayesian Networks*, IEEE Transactions on Knowledge and Data Engineering, IEEE Computer Society, 2007, 19, 1420-1432

[8] Mendes, E. The Use of a *Bayesian Network for Web Effort Estimation*, ICWE 2007, Web Engineering, 7th International Conference, Como, Italy, July 16-20, 2007, Proceedings, Springer, 2007, 4607, 90-104

[9] Kim, J. H., & Pearl, J. (1983) *A computational model for combined causal and diagnostic reasoning in inference systems*, In Proceedings of the eighth international joint conference on artificial intelligence, Karlsruhe (pp. 380-385)

[10] Saaty, T. *How to make a decision - The Analytic Hierarchy Process Interfaces*, Inst Management SCI, 1994, 24, 19-43 How to make a decision - The Analytic Hierarchy Process Interfaces, Inst Management SCI

[11] Miranda, E. *Improving subjective estimates using paired comparisons IEEE Software*, 2001, 18, 87-91