



Universidad Nacional Autónoma
de México

12
rej

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
ACATLAN

**' COMPARACION ENTRE LAS TECNICAS DE GRADUACION
DE WHITTAKER - HENDERSON Y LA GRADUACION
BAYESIANA, CON EJEMPLIFICACIONES.'**

**TESIS PROFESIONAL
QUE PARA OBTENER EL TITULO DE
A C T U A R I O
P R E S E N T A :
GULLERMINA GABRIELA LAMBARRI VAZQUEZ**

Asesor de Tesis: **MANUEL ORDORICA MELLADO**

ACATLAN, EDO. DE MEX. GENERACION 91-94



FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AL HOMBRE QUE MÁS RESPETO,
QUERO Y ADMIRO EN ESTE MUNDO...
¡MI PAPA!

A MI ABUEA CON TODO MI CARINO..

**A MI HERMANO POR GUARDAR
SILENCIO MIENTRAS ESTUDIABA..**

**Y POR SUPUESTO A MI DEBESINO (PAPI)
POR TODO EL APOYO QUE ME HA DADO..**

INDICE

Prólogo.	I
Introducción.	III
Capítulo I.	1
Nociones básicas de la graduación.	
1.1 Definición de graduación.	
1.2 Naturaleza de la graduación.	
1.2.1 Clasificaciones.	
1.3 Importancia de la graduación dentro del trabajo actuarial.	
Capítulo II	8
Consideraciones estadísticas, de probabilidad y de métodos numéricos.	
2.1 Problemas de estimación.	
2.2 Pruebas de suavización.	
2.3 Pruebas de ajuste.	
2.4 Conceptos de Métodos Numéricos.	
2.5 Conceptos de Estadística Bayesiana.	
Capítulo III	23
Fórmulas básicas de graduación y algunas de sus variantes.	
3.1 Determinación de la fórmula básica de Whittaker- Henderson.	
3.2 Variantes de la graduación de Whittaker-Henderson.	
3.3 Determinación de la fórmula básica de la graduación Bayesiana.	
3.4 Variantes de la graduación Bayesiana.	
Capítulo IV	45
Diferencias analíticas entre las dos técnicas.	
4.1 Diferencias en la concepción.	
4.2 Diferencias en las herramientas utilizadas.	
4.3 Ventajas y desventajas, similitudes y diferencias generales.	

Capítulo V	47
Aplicaciones numéricas.	
5.1 Tablas utilizadas.	
5.2 Graduación de las tablas por medio de las dos técnicas.	
5.3 Comparación de los resultados.	
Conclusiones	76
Bibliografía.	77
Anexos.	79

PROLOGO.

Uno de los temas poco recurridos por la comunidad actuarial de nuestra Escuela Nacional para la presentación de tesis es el concerniente a la construcción de tablas de mortalidad y por ende cualquier técnica que se encuentre relacionada con ella.

Uno de los temas de gran relación con la construcción de estas tablas es el trabajo de graduación que, como veremos, es el toque final en la construcción de cualquier tipo de tabla de decrementos. Al igual que cualquier materia, estas técnicas han ido desarrollándose con el paso del tiempo. En principio la graduación se llevaba a cabo por métodos gráficos, i.e., se graficaban los datos que se tenían y luego se iba corrigiendo la curva resultante hasta obtener una curva suave y continua. Luego se pasaron a métodos más sofisticados hasta llegar a aquellos que involucran un conocimiento más profundo de las Matemáticas, el Cálculo y la Estadística.

Por lo tanto una de las principales fuentes de motivación para realizar la presente tesis fue el de poder difundir un tema de tanta importancia en el campo actuarial y que tan pocas veces tomamos en cuenta. Recordemos que el cálculo actuarial no existiría sin las llamadas tablas de decrementos, llámense de mortalidad, de invalidez, rotación, etc., puesto que junto con los elementos financieros son el corazón de todo trabajo que presuma de llamarse actuarial.

Es interesante ver como el desarrollo de estas técnicas de graduación nos lleva a aplicar conocimientos de materias que muchas veces durante nuestra formación profesional como actuarios cuestionamos para que nos servirán. Fue realmente satisfactorio encontrar que al desarrollar el presente trabajo de tesis todos esos conocimientos adquiridos sobre Cálculo Diferencial e Integral, Estadística, Métodos Numéricos y Álgebra Lineal entre otros, se hayan fusionado de tal forma que sin su ayuda hubiera sido imposible fundamentar y desarrollar técnicas tan características de nuestra profesión como lo es la creación de tablas de decrementos.

También la que suscribe el presente trabajo exhorta a cualquier otro estudiante que ande buscando tema de tesis y que se interese por este tipo de cosas a que complemente y/o corrija este trabajo con la construcción de una tabla de mortalidad para hombres y otra para mujeres *confiables* que aún a la fecha en nuestro país no existen. Se hace la distinción entre mujeres y hombres porque recordemos que el comportamiento de la mortalidad entre los dos sexos es

diferentes y que esta diferencia *debe* ser tomada en cuenta para que así los cálculos actuariales sean más representativos que al utilizar una sola tabla para ambos sexos.

Por último la que ha realizado el presente trabajo está consciente de que existen fallas en el ejemplo práctico pero también espera que en un futuro esto sea mejorado y corregido por alguna exposición posterior.

G.G.L.V.
Marzo, 1995.

INTRODUCCION.

La Graduación tiene por tarea el convertir en curvas suaves y continuas las curvas de eventos en los que firmemente creemos que existe una interrelación entre sus vecindades y que no lo representan debido a que aun se observen cambios bruscos que no concuerden con nuestra creencia. Ejemplo claro, y por demás recurrido en esta tesis, es el caso de la mortalidad donde la interrelación consiste en un aumento de la probabilidad de muerte según aumente la edad del individuo (con excepción de los primeros años de vida en que la mortalidad suele ser más alta que en edades posteriores), es por eso que se busca que la curva se suavice y no presente picos que sugieran cambios muy drásticos de probabilidades de muerte de edad a edad. Además la curva debe ser continua dado que aunque la costumbre sea la de representar en las tablas de mortalidad sólo las edades enteras no se puede negar que este fenómeno puede presentarse en cualquier momento de la vida de un individuo.

Para lograr esto se recurre a las llamadas técnicas de graduación. Para nuestros propósitos utilizaremos dos de ellas (Whittaker-Henderson y Bayesiana), lo cual no significa que sean las únicas.

Entonces:

1. El objetivo del presente trabajo es el de realizar una comparación entre las técnicas de graduación de Whittaker-Henderson y la graduación de Bayes con la finalidad de poder aplicar algún tipo de criterio cuando se esté a punto de escoger entre alguna de ellas para poder llevarla a la práctica. Esta decisión deberá basarse en las ventajas que ofrezca cada técnica con respecto a la otra.

2. Existe otro objetivo que se pretende alcanzar y que es el demostrar cuan importante es el graduar la información que se utilizará como tablas de decrementos en los cálculos actuariales ya que si no se realiza dicha graduación esto causará que los cálculos obtenidos luzcan poco consistentes entre si.

3. La hipótesis existente antes de concluir el trabajo era la siguiente:

La técnica de Whittaker-Henderson es la más sencilla de aplicar por las herramientas de cálculo utilizadas, sin embargo los fundamentos teóricos de la graduación de Bayes son más atractivos puesto que involucran la utilización de información ya existente que pueda servir de respaldo.

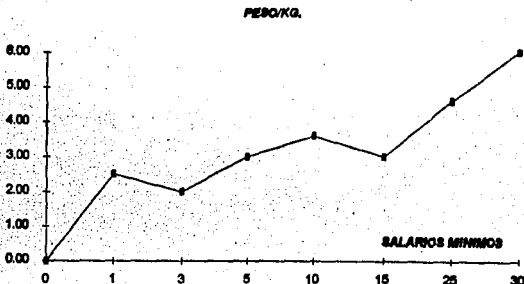
CAPITULO UNO. NOCIONES BASICAS DE LA GRADUACION.

1.1 Definición de graduación.

Antes que nada es de suma importancia comprender a que nos estamos refiriendo cuando hablamos de graduación y en que casos, y que cosas, podemos graduar.

No toda la información es susceptible de graduarse.

Se puede graduar toda aquella información en la que después de haberse calculado estimaciones iniciales se sospeche que existe una interrelación entre las vecindades de las estimaciones que generalmente representan *secuencias* de datos, por ejemplo, al medir el grado de desnutrición en niños de diferentes estratos económicos entre cero y un año de edad, se espera, en general, que entre mayor sea la solvencia económica de los padres el grado de desnutrición vaya disminuyendo. De acuerdo con esto supongamos que nuestras estimaciones iniciales al ser graficadas muestren el siguiente aspecto:



Por supuesto se debe tener en cuenta que lo aquí se presenta se trata solamente de un ejemplo ficticio que nos ayudará a comprender lo que estamos explicando.

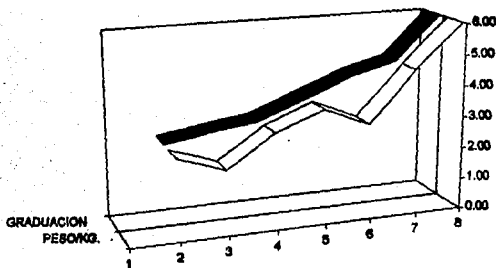
Observamos que la gráfica presenta picos, esto contradice a lo que intuitivamente proponíamos de que a mayor solvencia económica menor el grado de desnutrición, es decir, nuestra lógica nos indica que la gráfica debería ser la de

una función monótona creciente (sin saltos) y continua, lo que nos hace pensar que aún falta algo por hacer.

Esto que nos falta por hacer y que logrará una curva continua creciente es lo que llamamos *graduación de las estimaciones iniciales*.

El propósito de la graduación consiste, entonces, en obtener una curva representativa del fenómeno estudiado (en este caso desnutrición vs. solvencia económica) que sea una curva suave y lo más ajustada posible a las estimaciones originales (o iniciales).

Es decir, algo como lo siguiente:



Aquí podemos observar la curva graduada (curva gris) que está más acorde a la idea básica de interrelación entre los datos (nutrición creciente a mayor salario) que teníamos, es decir se suavizó la información a la vez que se cuidó de que la nueva curva tuviera un grado de ajuste a la original bastante bueno.

Llegado a este punto podemos ver que existen ciertas nociones básicas que deben ser tomadas en cuenta al realizar la graduación de estimaciones iniciales:

1.- Por supuesto lo primero a hacerse deberá ser la obtención de una representación inicial del fenómeno en estudio (estimaciones iniciales), lo cual se puede hacer con técnicas estadísticas ya conocidas.

2.- Mejorar la representación de nuestro primer modelo por medio de:

a) suavización,

b)ajuste,

que es precisamente lo que llamamos graduación y que es lo que se desarrollará en el presente trabajo.

Es necesario aclarar en este punto que no se están juzgando las técnicas estadísticas conocidas al querer "revisar" las estimaciones que estas nos proporcionan, es muy válido en estos momentos preguntarse ¿porqué revisar algo que se supone ha sido obtenido por métodos que ya han sido estudiados y aceptados?

La respuesta a esta pregunta la encontramos precisamente en la naturaleza del conjunto de datos a graduar en el que intuitivamente creemos que existe una interrelación entre sus elementos que no ha sido incluida por los simples procedimientos estadísticos conocidos.

Como ya se había mencionado anteriormente no toda información (o estimaciones iniciales) es susceptible de graduarse, en estos casos, lo calculado con los métodos estadísticos conocidos será más que suficiente.

Por ejemplo:

Podría hacerse una estimación del cambio en la tasa de natalidad, en un año, de los países latinoamericanos ordenándolos en forma alfabética. Sin embargo el pensar que los países en los que su nombre comience con la n -ésima letra del alfabeto tendrán una mayor (o menor) natalidad que aquellos que comiencen con la $(n-1)$ -ésima letra del alfabeto es algo que desde un principio resulta ilógico.

En este ejemplo se puede apreciar como no existe una interrelación válida entre los elementos de la secuencia formada, por lo cual no se puede justificar la creación de una curva continua monótonamente creciente o decreciente ajustada a las estimaciones iniciales. Es decir aquí la estadística ha hecho lo mejor que es posible para representar el fenómeno.

En general el incluir al tiempo como variable en estudio nos conduciría a la creación de una secuencia de estimaciones dignas de graduarse, es decir en el ejemplo anterior se pudo haber dicho que en un solo país en un periodo de cinco años la natalidad ha tendido a disminuir por lo que la curva que nos represente la natalidad en ese periodo en ese país deberá ser por lógica una curva decreciente, claro de antemano hemos *desechado* la idea de algún *suceso extraordinario* que

nos pudiera provocar un aumento considerable de natalidad cuando en el año inmediato anterior ésta había disminuido.

En base a todo lo anterior podemos ya construir la siguiente definición de acuerdo al criterio de la que suscribe la presente tesis:

Definición:

Graduar es el *arte* y la *técnica* de revisar estimaciones previamente calculadas para un fenómeno dado, de tal forma que se incluyan las creencias lógicas de interrelación (habiendo desechado de antemano sucesos extraordinarios debido a su *insignificancia*) entre los elementos de la muestra estudiada y así formar una curva *suave* y lo más *ajustada* posible a la curva *original*.

Se define a la graduación como un *arte* puesto que en gran medida la creación de la *curva graduada*, dependerá de la habilidad y destreza de quien realice el trabajo.

Se le define como una *técnica* puesto que ya existen procedimientos específicos (como los que más adelante se tratarán) para llevarla a cabo.

Para concluir esta sección diremos que hay que notar que graduamos simplemente lo que hemos observado (muestra) y suponemos que el fenómeno se mantendrá constante, si queremos crear un modelo que tome en cuenta los cambios a los que se sujeta el fenómeno en cuestión a través del tiempo¹, estaremos entrando a una tercer fase en la construcción de nuestro modelo llamado *proyección*. Este es un tema de gran utilidad pero que no desarrollaremos en este trabajo, sin embargo se hará mención de un ejemplo que nos ilustrará la importancia de la proyección, en el caso específico de proyecciones de población en Demografía:

"...Las proyecciones de población son importantes en la programación económica y social e indispensables para calcular los requerimientos futuros en materia de educación, empleo, vivienda, salud, seguridad social. Dichas proyecciones ayudan también a entender mejor la dinámica de los fenómenos demográficos, pues permiten estudiar los efectos de las variaciones de las tasas de natalidad, mortalidad y migración en la estructura por edades, así como predecir el futuro y comprender el pasado. Por su parte, la proyección

¹Principalmente el tiempo, puesto que el tiempo conlleva cambios en la tecnología, actitudes, políticas, economía, nuevos descubrimientos, etc.

de escenarios deseables permite proponer alternativas a la evolución de fenómenos demográficos que pueden provocar efectos negativos en el desarrollo del país.²

1.2 Naturaleza de la graduación.

Es nuestra intención demostrar que la graduación tiene una naturaleza estadística para así poder utilizar las herramientas que esta ciencia nos ofrece.

Esta naturaleza estadística es clara si tomamos en cuenta que estamos tratando de *estimar* lo más correctamente posible las verdaderas tasas de cambio en una serie de observaciones con relación a algún fenómeno dado que suelen ser muy irregulares de acuerdo a nuestra lógica. Al llevar a cabo dicha estimación se tienen que aplicar *herramientas matemáticas y técnicas estadísticas*, sin las cuales no podría llevarse a cabo la graduación. Comprendiendo esto será más fácil el entender el porqué de la utilización de conceptos tales como variable aleatoria, media, varianza, intervalos de confianza, etc., en las diferentes técnicas de graduación.

La única diferencia es que la estimación estadística se basa exclusivamente en los datos observados y ahí finaliza su tarea, mientras que la graduación toma las estimaciones iniciales conjuntamente con esa creencia que se tiene de interrelación entre las vecindades de las estimaciones y a partir de ellas realizar su trabajo revisándolas y convirtiéndolas en estimaciones finales o graduadas.

1.2.1 Clasificaciones.

La graduación se divide principalmente en dos tipos de métodos:

- 1.- Graduación de estimaciones iniciales en forma tabular (tablas).
- 2.- Graduación de estimaciones iniciales representadas por una función, es decir en forma paramétrica.

Un buen ejemplo del primer caso son las tablas de mortalidad, invalidez, rotación, y en general las tablas de decrementos tan utilizadas en el trabajo actuarial.

²Ordorica Mellado, Manuel. *La población de México en los albores del siglo XXI: ¿predicción o proyección?* Comercio Exterior. México, julio de 1993. p. 634.

Para el segundo caso tenemos aquellas funciones clásicas representativas de la mortalidad tales como el modelo de Gompertz:

$$l_x = kg^{c^x}$$

o bien el modelo de Makeham:

$$l_x = ks^{-x}g^{c^x}$$

que como bien es sabido representan apropiadamente la típica curva de l_x sobre un considerable rango de edades.

En el presente trabajo se tratará con técnicas de graduación para estimaciones iniciales en forma tabular (primer caso) dado que es más común encontrarlas de esta forma que representadas por medio de una función. Dichas técnicas son:

- Graduación de Whittaker-Henderson.
- Graduación de Bayes.

1.3 Importancia de la graduación dentro del trabajo actuarial.

Se comenzará esta sección haciendo una cita:

"El actuario espera poder utilizar sus tablas de mortalidad para hacer cálculos de primas, reservas, anualidades y demás. No existe ganancia alguna en el suponer algo distinto al hecho de que la mortalidad varía en forma regular y continuamente. Irregularidades caprichosas en las tablas de edad a edad interfieren el orden de progresión de las primas, etc., que resulta inconsistente con la lógica común de que tales cosas deberían ser regulares, causando así una desconfianza bastante justificable."³

Como sabemos una de las tareas principales del actuario es la de realizar la construcción de modelos de sobrevivencia generalmente mostrados en forma tabular (tablas de mortalidad). Estos modelos son necesarios pues sin ellos no

³Miller. Comentario tomado de: London, D. *Graduation. The revision of estimates*. Winsted, Connecticut:ACTEX publications, 1985.

podría ser posible el cálculo de primas para los seguros o las aportaciones a un fondo de un plan de pensiones, por citar algunos ejemplos.

Dado que el actuario trata con eventos que regularmente forman secuencias hemos llegado a la conclusión de que el material utilizado por él es susceptible de graduarse, un ejemplo obvio es la interrelación que guardan las tasas de mortalidad a diferentes edades, éstas deben ir aumentando conforme lo haga la edad del individuo (con excepción de los primeros años de vida en que la mortalidad es mayor que en los subsecuentes) y esto a su vez se verá reflejado en los cálculos actuariales que se realicen. Por ejemplo pensemos en un seguro de vida, de una u otra forma intuitivamente todos creemos que a una persona con mayor edad que otra deberá cobrarsele una prima más alta por un beneficio del mismo monto, dado que para ella se cuenta con un menor tiempo (debido a su mayor probabilidad de morir en un tiempo dado por ser más "vieja") para crear el fondo que le otorgará dicho beneficio. Generalmente esta interrelación de la que hablamos no es tomada en cuenta al calcular las estimaciones iniciales (que son hechas independientemente por cada edad) provenientes de los censos o de estadísticas de las empresas, es por eso que tenemos, repito, que graduar todo este trabajo inicial.

Aclarado este punto mencionaremos que de aquí en adelante la información a la que se haga referencia en la graduación será principalmente aquella utilizada por los actuarios, tales como la mortalidad, invalidez, rotación, etc.

CAPITULO DOS.
CONSIDERACIONES ESTADISTICAS, DE PROBABILIDAD Y
DE METODOS NUMERICOS.

*Notación utilizada...

Siendo el libro "Graduation..." de Dick London uno de los principales pilares de la presente tesis, se ha decidido adoptar el tipo de notación por él utilizada, que a continuación será descrita:

<i>Símbolo</i>	<i>Significado</i>
x	El índice que define a la secuencia.
t_x	La secuencia de valores reales que serán estimados.
U_x	Las variables aleatorias estimadores de t_x , asumidas como proporciones binomiales insesgadas.
n_x	Los tamaños de muestra (expuestos) en los modelos binomiales asumidos.
u_x	Las realizaciones particulares de U_x , nuestras estimaciones iniciales (o valores observados) de t_x .
v_x	Las estimaciones revisadas de t_x , nuestros valores graduados.
V_x	La variable aleatoria de la cual v_x es una realización particular.
w_x	La secuencia de ponderaciones usadas en una graduación.
E_x	La variable aleatoria de estimación del error incurrido.
e_x	Las realizaciones particulares de E_x .

2.1 Problemas de estimación.

2.1.1 Analogía de un proceso Binomial con el comportamiento de la mortalidad.

Primeramente hay que tener en cuenta que para graduar existen tres cosas:

1. Las t_x que son los *valores reales* que desconocemos pero que queremos estimar.
2. Las u_x que son las *estimaciones iniciales* con que contamos, y
3. Las v_x , *valores graduados*, que pretenden ser nuestra mejor representación posible de las t_x .

Lo principal al hacer una estimación inicial de las t_x es escoger el modelo que mejor represente el comportamiento de la población dado cierto evento, e.g. mortalidad, en este caso las t_x representarán la probabilidad de muerte de un individuo con edad x .

El evento mortalidad, como es bien sabido, se presenta en forma dicótoma, te encuentras muerto o no.

Cuando hablamos de eventos que sólo tienen dos posibles resultados, éxito o fracaso, hablamos de procesos Bernoulli (un solo ensayo) o de procesos Binomiales (varios ensayos).

Recordemos que los ensayos dentro de un modelo Binomial son independientes entre sí y aunque la mortalidad no se distribuye exactamente como una Binomial ya que los ensayos no son independientes (el hecho de que un individuo sobreviva al próximo año está condicionado a que sobreviva este), podemos considerarlo como tal ya que regularmente el número de expuestos en la muestra es lo suficientemente grande como para que el error que se produce sea significativo.

Para desarrollar el modelo Binomial correspondiente supongamos que el resultado aleatorio de un volado (cara o cruz) representará al evento mortalidad (sobrevivir o no).

Entonces:

Sea t la probabilidad de obtener un cierto número de caras en n volados, para poder estimar esta t lo que podemos hacer es observar la *proporción de ocurrencias*, llamémosle μ , al realizar estos n volados.

Si definimos a H como la variable aleatoria que represente el *número de caras obtenidas en n ensayos* tenemos que:

$$E[H] = nt \text{ y } Var(H) = nt(1-t)$$

Ahora si definimos a la variable aleatoria U como la *proporción de caras obtenidas en n volados*, obtendremos:

$$U = \frac{H}{n}$$

que es lo que conocemos como una *proporción binomial*, por razones obvias.

Entonces:

$$E[U] = \frac{1}{n} E[H] = t$$

y

$$Var(U) = \frac{1}{n^2} \cdot Var(H) = \frac{t(1-t)}{n}$$

Hay que recordar que μ es una estimación de t (i.e. $\mu = t$).

Dado que U nos será de gran utilidad a lo largo de este trabajo, es de interés remarcar las siguientes características:

1. U es un estimador insesgado de t , ($E[U] = t$).

2. U es el estimador de máxima verosimilitud⁴ de t .
3. U es también el estimador del método de momentos⁵.
4. Notemos que $Var(U) = \frac{t(1-t)}{n}$, lo que significa que la varianza de una proporción Binomial es *inversamente proporcional* al tamaño de la muestra, propiedad de suma importancia en la graduación.
5. Si n es lo suficientemente grande, la distribución Binomial de U puede ser aproximada mediante una distribución normal⁶.

Dado que frecuentemente necesitaremos un valor numérico para la varianza y viendo que esta está ligada al valor desconocido de t utilizaremos en su lugar:

$$\hat{Var}(U) = \frac{u(1-u)}{n}$$

lo cual se justifica porque $u = \hat{t}$.

2.2 Pruebas de suavización.

Esencialmente el obtener un indicador del grado de suavización en una curva graduada es meramente relativo. Lo más fácil es graficar los datos graduados y observar que tan suave es la curva obtenida. Por otro lado si lo que queremos es una medida *numérica*, podemos auxiliarnos de las diferencias progresivas finitas para comparar el grado relativo de suavización en diferentes graduaciones de la *misma* información y más aún, de la misma información sin graduar. Este número índice (útil sólo para fines comparativos, i.e. sin significado por si solo) puede ser obtenido como la suma de los cuadrados (o valores absolutos) de dichas diferencias.

$$\sum (\Delta^2 v_i)^2$$

⁴Ver Anexo Uno.

⁵Ver Anexo Uno.

⁶Ver Anexo Dos.

Por ejemplo, si pensamos que nuestra información es una curva de naturaleza cúbica lo más lógico es considerar diferencias cuartas ($z=4$) para calcular su suavización.

2.3 Pruebas de ajuste.

"Siempre se ha discutido que los valores graduados no deberían alejarse demasiado de los valores iniciales; esto es, *debe* existir algún grado de *ajuste* entre los dos."

Una forma de asegurar un buen ajuste de los datos graduados a los datos iniciales es tomando el número de expuestos n_x (muestra) lo suficientemente grande (ley de los grandes números, proporción binomial) lo cual implicaría un mejor ajuste, es decir desviaciones más pequeñas, en comparación con aquellas muestras de tamaño más pequeño con mayor desviación.

Consideremos las siguientes medidas numéricas de ajuste:

$$F_1 = \sum_x w_x (v_x - u_x)$$

y

$$F_2 = \sum_x w_x (v_x - u_x)^2$$

La primera tiene una explicación lógica, ponderar las diferencias entre los valores graduados y los que observamos, sin embargo F_1 tiende a cancelarse por lo cual es mejor utilizar F_2 que al elevar al cuadrado las diferencias nos evita los problemas con los signos.

Por lo tanto para juzgar esta medida numérica de ajuste tomemos en cuenta que si U_x es aproximadamente normal, con:

$$E[U_x] = t_x \quad \text{y} \quad \text{Var}(U_x) = \frac{t_x(1-t_x)}{n_x}$$

⁷London, D. *Graduation. The revision of estimates*. Winsted, Connecticut: ACTEX publications, 1985. p. 16.

entonces:

$$\frac{(U_x - t_x)^2}{\text{Var}(U_x)}$$

tiene una distribución χ^2 con 1 grado de libertad. Más aún, si cada una de las U_x es independiente, tenemos que:

$$F = \sum_x \frac{(U_x - t_x)^2}{\text{Var}(U_x)}$$

es una variable aleatoria con una distribución χ^2 con n grados de libertad y n términos en la suma, además sabemos que:

$$E[F] = E[\chi_{(n)}^2] = n$$

Por supuesto que no conocemos las t_x pero podemos utilizar las v_x que son nuestras mejores representaciones de las t_x , para poder calcular

$$F_2 = \sum_x w_x (v_x - u_x)^2 \text{ por medio de } w_x = \frac{n_x}{v_x(1-v_x)} \cong \frac{n_x}{t_x(1-t_x)} \cong \frac{1}{\text{Var}(U_x)}^8.$$

2.4 Conceptos de Métodos Numéricos.

En el campo actuarial constantemente es necesario obtener respuestas numéricas a problemas físicos (e.g. representación de la mortalidad). El primer paso para resolver dichos problemas es el transformarlos en problemas matemáticos. Así tenemos por ejemplo la aproximación de Woolhouse para poder calcular anualidades pagaderas m veces al año que nos dice:

⁸Ver página 10.

$$\frac{1}{mD_x} \sum_{t=1}^{\infty} D_{x+t/m} = \frac{1}{D_x} \left[\sum_{t=1}^{\infty} D_{x+t} + \frac{m-1}{2m} D_x + \frac{m^2-1}{12m^2} \cdot \frac{dD_x}{dx} \dots \right]$$

donde:

m = frecuencia de pagos en el año

D_x = Valor conmutado

y que es a su vez la representación numérica o mejor dicho el *método numérico* en el cual hemos transformado nuestro problema de cálculo de anualidades contingentes pagaderas m veces al año.

Muchas veces el encontrar soluciones numéricas a problemas físicos depende de un escenario de ensayo y error hasta acercarnos a lo que buscamos o a lo que nuestra intuición nos dicte.

La interpolación cae dentro del campo de los métodos numéricos, y en nuestro caso nos será de utilidad, específicamente para la graduación de Whittaker-Henderson, ya que las técnicas de graduación que se presentan son para estimaciones representadas en *forma tabular*.

Recordemos que llamamos interpolación al proceso de pasar una curva a través de puntos dados (tabla) para determinar valores funcionales $f(x)$ para valores de x que no se muestren explícitamente en la tabla.

Particularmente, dentro del tema de la Interpolación, haremos un recorrido por el método de diferencias finitas *progresivas*. Pero primeramente recordemos la fórmula de interpolación de Gregory-Newton, para la construcción de un polinomio, que nos dice:

$$p(u) = f(x_1) + \frac{\Delta f(x_1)}{1!} u + \frac{\Delta^2 f(x_1)}{2!} u(u-1) + \frac{\Delta^3 f(x_1)}{3!} u(u-1)(u-2) + \dots$$

$$+ \frac{\Delta^n f(x_1)}{n!} u(u-1)(\dots)[u-(n-1)]$$

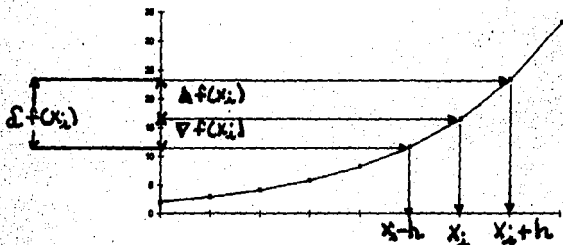
Diferencias finitas.

Una diferencia finita (o simplemente diferencia) de la función $f(x)$ es el valor de la función en un punto x_1 menos el valor de un segundo punto x_2 .

Algebraicamente, esto se representa como:

$$f(x_1) - f(x_2)$$

Geoméricamente las diferencias son representadas por intervalos a lo largo del eje vertical de un plano cartesiano:



Las cantidades $\Delta f(x_i)$, $\nabla f(x_i)$ y $\delta f(x_i)$ son llamadas todas diferencias.

En particular la que nos interesa $\Delta f(x_i)$ es llamada diferencia progresiva.

En una tabla, las diferencias se van obteniendo por sustracción de valores sucesivos, i.e., $f(x_{i+1}) - f(x_i)$ y se les denomina diferencias primeras, segundas, etc. según sea la primera resta, la segunda, etc., e.g.

x	$f(x)$	Diferencia primera	Diferencia segunda	Diferencia tercera
1	4	17		
2	21	27	10	
3	48	37	10	0
4	85			

Entonces las diferencias progresivas de esta tabla, se obtienen escogiendo un punto inicial x_i y trazando una diagonal en forma descendiente como se ilustra en dicha tabla.

Generalizando:

x_i	$f(x_i)$	$\Delta f(x_i)$	$\Delta^2 f(x_i)$	$\Delta^3 f(x_i)$	$\Delta^4 f(x_i)$
x_{i+h}	$f(x_{i+h})$	$\Delta f(x_{i+h})$	$\Delta^2 f(x_{i+h})$	$\Delta^3 f(x_{i+h})$	
x_{i+2h}	$f(x_{i+2h})$	$\Delta f(x_{i+2h})$	$\Delta^2 f(x_{i+2h})$	$\Delta^3 f(x_{i+2h})$	
x_{i+3h}	$f(x_{i+3h})$	$\Delta f(x_{i+3h})$	$\Delta^2 f(x_{i+3h})$		
x_{i+4h}	$f(x_{i+4h})$				

Expresado algebraicamente:

$$\Delta f(x_i) = f(x_{i+h}) - f(x_i)$$

$$\begin{aligned}\Delta^2 f(x_i) &= \Delta f(x_{i+h}) - \Delta f(x_i) \\ &= [f(x_{i+2h}) - f(x_{i+h})] - [f(x_{i+h}) - f(x_i)] \\ &= f(x_{i+2h}) - 2f(x_{i+h}) + f(x_i)\end{aligned}$$

$$\begin{aligned}\Delta^3 f(x_i) &= \Delta^2 f(x_{i+h}) - \Delta^2 f(x_i) \\ &= [\Delta f(x_{i+2h}) - \Delta f(x_{i+h})] - [\Delta f(x_{i+h}) - \Delta f(x_i)] \\ &= \Delta f(x_{i+2h}) - 2\Delta f(x_{i+h}) + \Delta f(x_i) \\ &= [f(x_{i+3h}) - f(x_{i+2h})] - 2[f(x_{i+2h}) - f(x_{i+h})] + [f(x_{i+h}) - f(x_i)] \\ &= f(x_{i+3h}) - 3f(x_{i+2h}) + 3f(x_{i+h}) - f(x_i)\end{aligned}$$

etc., etc...

Observemos que los coeficientes de la función en $\Delta^k f(x_i)$ son los mismos que aquellos obtenidos por el binomio de Newton.

Por lo tanto, la k-ésima diferencia progresiva en el punto x_i es:

$$\Delta^k f(x_i) = f(x_{i+kh}) - \binom{k}{1} f(x_{i+(k-1)h}) + \binom{k}{2} f(x_{i+(k-2)h}) + \dots + (-1)^k f(x_i)$$

Si hacemos $i=1$ entonces las diferencias progresivas corresponderán a los coeficientes de la fórmula de interpolación de Gregory-Newton.

Por lo tanto el polinomio que interpole los valores de la tabla será igual al que resulte de sustituir las diferencias progresivas en la fórmula de Gregory-Newton.

Las diferencias progresivas cuentan con las siguientes propiedades:

• La aplicación del operador Δ a una función polinomial reduce el grado del polinomio en uno, i.e., si $f(x)$ es un polinomio de grado n , $\Delta f(x)$ es un polinomio de grado $n-1$, de lo cual se sigue que si $f(x)$ es un polinomio de grado n , $\Delta^n f(x)$ es un polinomio de grado cero, o mejor dicho una constante;

$$\Delta^n f(x) = c$$

y

$$\Delta^{n+1} f(x) = 0$$

• El operador de diferencias progresivas es distributivo:

$$\Delta[f(x) + g(x)] = \Delta f(x) + \Delta g(x)$$

• Por último, también es conmutativo con respecto a una constante:

$$\Delta[c \cdot f(x)] = c \cdot \Delta f(x)$$

2.5 Conceptos de Estadística Bayesiana.

La inferencia Bayesiana tiene por principio utilizar la información histórica existente acerca de los valores que queremos estimar de una población para obtener estimaciones actuales de nuestros valores reales desconocidos t_x ; i.e., nuestras estimaciones *a posteriori* se encuentran *condicionadas* a los valores estimados *a priori* con los que contamos. Sin embargo si no contamos con observaciones *históricas*, entonces es posible aplicar una distribución que a nuestro juicio se ajuste a nuestros valores reales t_x .

Este tipo de estadística toma su nombre precisamente del *teorema de Bayes* que es su eje central y fundamental.

El teorema de Bayes nos dice:

$$f_{TV}(t|u) = \frac{f_{U|T}(u|t) \cdot f_T(t)}{f_U(u)}$$

donde T y U son vectores de v.a.i.i.d.

Por lo que $f_T(t)$ representa nuestra distribución *a priori*, $f_{U|T}(u|t)$ como se comportaron los valores observados conociendo el comportamiento de los valores reales, distribución *a posteriori*, y $f_U(u)$ no es mas que una distribución marginal obtenida a partir de $f_{U|T}(u|t)$.

Para facilitar la aplicación de este teorema es práctica común auxiliarse de las llamadas distribuciones conjugadas.

Las distribuciones conjugadas se definen de la siguiente manera:

"Si la función de probabilidad anterior, $f_T(t)$, es un miembro de una familia de distribuciones tal que sea *conjugada* para el modelo del experimento $f_{U|T}(u|t)$, entonces la función de probabilidad posterior, $f_{T|U}(t/u)$, será un miembro de esa misma familia, y por consecuencia, poseerá las mismas propiedades que posea la función de probabilidad anterior."⁹

Es decir si al multiplicar $f_T(t)$ por $f_{U|T}(u|t)$ y dividirla entre $f_U(u)$, según el teorema de Bayes, obtenemos una $f_{T|U}(t/u)$ que pertenezca a la misma familia de distribuciones que $f_T(t)$ significará que $f_T(t)$ es la distribución conjugada para $f_{U|T}(u|t)$.

Ejemplos de distribuciones conjugadas son:

⁹London, D. *Graduation. The revision of estimates*. Winsted, Connecticut: ACTEX publications, 1985. p. 73.

**DISTRIBUCION DE LA
MUESTRA $f_{U|T}(u|t)$.**

- 1 Binomial
- 2 Binomial Negativa
- 3 Multinomial
- 4 Poisson
- 5 Exponencial con media λ^{-1}
- 6 Normal con varianza conocida,
pero con media desconocida
- 7 Normal con media conocida
pero varianza desconocida
- 8 Multinormal

**DISTRIBUCION
CONJUGADA $f_T(t)$.**

- Beta.
- Beta.
- Dirichlet.
- La media es Gamma.
 λ es Gamma.
- La media es Normal.
- La varianza es una
Gamma Invertida.
- Multinormal

Por ejemplo:

Si $f_T(t)$ es Beta y $f_{U|T}(u|t)$ es Binomial o Binomial Negativa, entonces

$f_{T|U}(t|u)$ será Beta.

O si $f_T(t)$ es Multinormal y $f_{U|T}(u|t)$ es también Multinormal, entonces

$f_{T|U}(t|u)$ será Multinormal.

La siguiente ilustración nos ayudará a comprender mejor esta metodología:

Ilustración.

Supongamos que estamos buscando el parámetro binomial t que nos represente la tasa verdadera de caras en el experimento de realizar un volado. En un proceso formal Bayesiano debemos expresar nuestra creencia inicial acerca de t por medio de una distribución de probabilidad más que por medio de un valor específico de t .

Entonces, sea T una v.a. con distribución Beta, i.e.,

$$f_T(t) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} t^{a-1}(1-t)^{b-1}, \quad 0 \leq t \leq 1,$$

donde $a > 0$ y $b > 0$ son parámetros. Fácilmente se comprueba que

$$\int_0^1 f_T(t) dt = 1 \quad \text{dado que} \quad \int_0^1 t^{a-1}(1-t)^{b-1} dt \quad \text{es la función Beta,}$$

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Podemos apreciar que el experimento (volados) es claramente binomial, entonces, si H es la variable aleatoria para el número de caras en n volados, tendremos que:

$$p_{H|T}(h|t) = \frac{n!}{h!(n-h)!} t^h (1-t)^{n-h}, \quad 0 \leq t \leq 1, \quad h = 0, 1, \dots, n.$$

La distribución marginal $p_H(h)$ se encuentra mediante:

$$p_H(h) = \int_0^1 p_{H|T}(h|t) \cdot f_T(t) dt,$$

lo cual da por resultado:

$$\frac{\Gamma(a+b)n!}{\Gamma(a)\Gamma(b)h!(n-h)!} \cdot \frac{\Gamma(a+b)\Gamma(b+n-h)}{\Gamma(a+b+n)}.$$

Aplicando el teorema de Bayes encontraremos que la función de distribución posterior es igual a:

$$f_{\pi_H}(t|h) = \frac{\Gamma(a+b+n)}{\Gamma(a+h) \cdot \Gamma(b+n-h)} t^{a+h-1} (1-t)^{b+n-h-1}$$

la cual es nuevamente una distribución Beta, donde a es reemplazada por $a+h$ y b por $b+n-h$. Por lo tanto hemos comprobado que la distribución Beta es conjugada a la Binomial, lo cual fue parte de la motivación para escogerla como nuestra distribución *a priori*.

Por último, recordando que lo que queríamos era un valor para t podemos tomar la media de la distribución posterior que hemos obtenido para t dados los resultados de la v.a. H , puesto que la media nos representa el valor que en promedio siempre tenderemos a obtener.

CAPITULO TRES.
FORMULAS BASICAS DE GRADUACION Y ALGUNAS DE
SUS VARIANTES.

3.1 Determinación de la fórmula básica de Whittaker-Henderson.

Fórmula básica.

Básicamente la fórmula de graduación de Whittaker-Henderson es una combinación lineal de las dos principales características con que debe contar toda técnica de graduación:

- Ajuste.-

$$F = \sum_{x=1}^n w_x (v_x - u_x)^2$$

- Suavización.-

$$S = \sum_{x=1}^{n-x} (\Delta^2 v_x)^2$$

F representa la dispersión de los datos y S representa la curva polinomial a la que se piensa que el comportamiento de los datos pudiera parecerse.

La combinación lineal es llevada a cabo de la siguiente forma:

$$M = F + hS$$

$$M = \sum_{x=1}^n w_x (v_x - u_x)^2 + h \sum_{x=1}^{n-x} (\Delta^2 v_x)^2$$

A continuación se irá desglosando y estudiando cada una de las componentes de esta fórmula, con la intención de ir comprendiéndola en su estructura.

Observemos primeramente su forma general:

$$M = F + hS$$

si recordamos el modelo de regresión lineal simple:

$$y = a + bx + e_x$$

e interpretamos M utilizando la comparación de acuerdo a esto, tenemos que el valor de M tiene un valor inicial (F , datos sin graduar) aún sin agregarle ningún otro tipo de elemento. El otro elemento que en este caso modifica a M es precisamente el grado de suavización, o semejanza a algún polinomio deseado S y por último tenemos un elemento de error e_x , que en el caso de la fórmula de Whittaker-Henderson es igual a cero, suponiendo que este error se distribuye normalmente con parámetros $(0, \sigma^2)$.

Por lo tanto, la función de h en esta fórmula es ni más ni menos que la de un coeficiente de regresión.

Es decir, la h nos determina el balance entre el mayor ajuste posible (menor valor posible de F) y la mayor suavización posible (menor valor posible de S). O en otras palabras; es bien sabido que S representa un polinomio de grado $Z-1$, entonces la magnitud de h determina que tan cercanos estarán los valores graduados al polinomio. Mientras mayor sea la magnitud de h , más se parecerán los datos graduados al polinomio; entre menor sea la magnitud de h más se parecerán los datos graduados a sus valores originales.

¿Cómo seleccionar la constante h de suavización?

Una de las formas de hacerlo es mediante ensayo y error, comparando los resultados para distintas pruebas con el mismo conjunto de datos y manteniendo en mente que si deseamos un mayor acercamiento de los datos graduados a los originales debe $h \rightarrow -\infty$ o viceversa, si queremos un mayor acercamiento de los datos graduados a la curva polinomial en cuestión, debe $h \rightarrow \infty$.

Regularmente se busca minimizar $F + hS$ tomando en consideración una h positiva. Si $h=0$, se enfatiza por completo el ajuste de los datos graduados a los originales, y en tanto $h \rightarrow \infty$ se irá dando más énfasis en el suavizamiento (semejanza al polinomio).

Es decir:

$$0 \leq h < \infty$$

lo cual es un rango bastante amplio. Pensemos entonces en hacer la siguiente simplificación:

$$(1-h)F + hS$$

$$\text{donde: } 0 \leq h \leq 1.$$

i.e., cuando $h=1$, existirá un completo énfasis en la suavización, y cuando $h=0$ existirá un completo énfasis en el ajuste. Esta modificación en la función a minimizar en el proceso de la graduación es equivalente a encontrar el balance apropiado tomando en consideración cierta ponderación del lado izquierdo, en lugar de tomarla solamente del lado derecho de la fórmula.

Más adelante regresaremos a esta cuestión, mientras tanto continuemos con nuestro análisis.

La parte de M que nos proporciona el ajuste a los datos originales:

$$F = \sum_{x=1}^n w_x (v_x - u_x)^2$$

será mejor entendida si mantenemos presente el concepto de varianza otorgada por la estadística.

Entendemos por varianza una medida del esparcimiento o dispersión de los valores que puede tomar la variable X correspondiente, con respecto a una *media representativa* \bar{X} . Esta medida es la diferencia básica entre dos muestras con la misma media, lógicamente al comparar dos muestras con la misma media se confiará más en la *representatividad* de la media de aquel conjunto en que la varianza sea menor.

Recordemos que la varianza muestral es igual a:

$$\text{Var} = \frac{\sum (\bar{X} - X_i)^2}{n-1}$$

Igualmente en el caso de la fórmula de Whittaker-Henderson, con F estamos midiendo el grado de ajuste (dispersión) de nuestros valores representativos v_x con respecto a los originales (muestra) u_x , y así de alguna forma poder representar a los valores reales t_x , la w_x en F hará entonces las veces de promedio (o ponderación) que obtenemos en la varianza al dividir entre $n-1$, y la podremos calcular por medio de:

$$w_x = \frac{n_x}{u_x(1-u_x)}$$

La suavización:

$$S = \sum_{x=1}^{n-f} (\Delta^x v_x)^2$$

Como ya se había mencionado S representa la curva polinomial a la que se piensa que el comportamiento de los datos pudiera parecerse.

Sabemos también que S son diferencias progresivas, la cuestión está en cómo obtener el orden de dichas diferencias que mejor nos representen nuestros datos muestrales.

Una sugerencia es la de manipular dichas diferencias hasta llegar a un nivel en donde sus valores tiendan todos a cero y tomar el valor de esa diferencia (diferencia tercera, diferencia cuarta, etc.) menos uno, ya que recuérdese que $\Delta^{n+1} f(x) = 0$ si y sólo si $f(x)$ es un polinomio de grado n , o bien equivalentemente tomar n de $\Delta^n f(x) = c$ donde c es una constante, i.e., donde veamos que las diferencias progresivas de los datos originales sean más o menos constantes.

Con esto ya hemos descrito todas las partes de la fórmula M .

Así tenemos que M puede ser cualquier valor dependiendo de los datos observados y aquellos valores graduados que escojamos, principalmente, pero lo más lógico es escoger los valores graduados de tal forma que estos difieran de los originales lo menos posible para así asegurar la mejor representación de los valores reales, es decir la diferencia:

$$v_x - u_x$$

tiene que ser mínima.

Al hacer esto a la par estaremos minimizando el valor de M por lo tanto para encontrar los valores de las v_x que lo minimicen, debemos derivar (parcialmente) a M , con respecto a cada una de las v_x (n en total) e igualar dichas derivadas a cero, lo que nos resultará en n ecuaciones con n incógnitas (las v_x), al resolver el sistema de ecuaciones resultante habremos encontrado los valores v_x que no son otra cosa más que los valores graduados buscados.

Ahora bien recordemos que proponíamos la siguiente simplificación:

$$(1 - N')F + N'S$$

$$\text{donde: } 0 \leq N' \leq 1.$$

Es entonces ésta, la expresión que derivaremos parcialmente. Para esto nos auxiliaremos del álgebra lineal, con el propósito de obtener una representación matricial de nuestro modelo y simplificar las operaciones.

Realización de la minimización de M .

M se minimiza de la siguiente manera:

$$\frac{\partial M}{\partial v_r} = 0, \text{ para } r=1,2,\dots,n$$

Como ya hablamos mencionado nos auxiliaremos del álgebra lineal, por lo tanto denotaremos por letras **negritas** a las representaciones tanto de vectores como de matrices.

Sean u y v los vectores de estimaciones iniciales y estimaciones revisadas (valores graduados) respectivamente. Utilizando y^t para denotar la transpuesta de y , tenemos que:

$$u^t = [u_1, u_2, \dots, u_n] \text{ y } v^t = [v_1, v_2, \dots, v_n]$$

Además sea:

$$\begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_n \end{bmatrix}$$

una matriz diagonal $n \times n$ que contiene las ponderaciones w_x . Sea k_z una matriz especial que contiene los coeficientes binomiales de orden z de las diferencias finitas progresivas, con su respectivo signo, tal que el producto $k_z v$ es el vector que contiene los valores de $\Delta^z v_x$. Por lo tanto si existen n valores v_x , tal que v sea un vector $n \times 1$, k_z será de dimensión $(n-z) \times n$ y $k_z v$ será un vector de $(n-z) \times 1$.

Por ejemplo:

Si $z=2$ y $n=6$, tendremos:

$$k_2 = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix}$$

Es de fácil verificación que $k_2 v$ será igual a $\Delta^2 v_x$. En cuanto a los signos, el primer elemento de $k_z (a_{11})$ será positivo si z es un número par, y negativo si z es un número impar.

Definidas así nuestras matrices v , u , w , y k_z , tenemos que la función M queda representada por:

$M = (v-u)^t w(v-u) + h(k_2 v)^t k_2 v$

(Recordando que M se encuentra de la forma;

$$\sum_{x=1}^n w_x (v_x - u_x)^2 + h \sum_{x=1}^n (\Delta^2 v_x)^2).$$

Si introducimos la modificación:

$$M = (1-h')F + h'S$$

nuestra representación matricial queda finalmente de la siguiente forma:

$$M = (1-h')(v-u)^t w(v-u) + h'(k_2 v)^t k_2 v$$

Ahora derivando parcialmente a M respecto a v tenemos:

$$\begin{aligned} \frac{\delta M}{\delta v} &= \frac{\delta}{\delta v} (1-h')(v-u)^t w(v-u) + h'(k_2 v)^t k_2 v \\ &= \frac{\delta}{\delta v} (1-h')(v-u)^t w(v-u) + h' v^t k_2^t k_2 v \\ &= 2(1-h') w(v-u) + 2h' k_2^t k_2 v \\ &= (2-2h')(wv - wu) + 2h' k_2^t k_2 v \\ &= 2wv - 2wu - 2h' wv + 2h' wu + 2h' k_2^t k_2 v \end{aligned}$$

Igualando todo lo anterior a cero, nos queda:

$$= 2wv - 2wu - 2h' wv + 2h' wu + 2h' k_2^t k_2 v = 0$$

$$2wv - 2h' wv + 2h' k_2^t k_2 v = 2wu - 2h' wu$$

$$wv - h' wv + h' k_2^t k_2 v = wu - h' wu$$

$$[(1-h')w + h'k^t k]v = (1-h)wu \dots (1)$$

y sustituyendo $c = [(1-h')w + h'k^t k]$, en (1), obtenemos:

$$cv = (I-h)wu$$

Estamos ahora en posibilidades de resolver el sistema de ecuaciones, que en forma matricial hemos obtenido, mediante algún método numérico como por ejemplo el método de eliminación de Gauss-Jordan, o métodos más sofisticados como lo es el método de factorización de Choleski¹⁰.

Dado que se puede demostrar que c es una matriz no singular se tiene como consecuencia que su inversa c^{-1} siempre existe y por lo tanto:

$$cv = (I-h)wu$$

siempre podrá ser resuelta por medio de:

$$v = (I-h)wuc^{-1}$$

con lo que finalmente habremos obtenido los valores graduados v_x .

Para simplificar el trabajo, (y eliminar la utilización del método de Choleski), se adoptó como lenguaje de programación el lenguaje APL dado que resulta bastante cómodo debido a su gran capacidad en el manejo de arreglos matriciales y por la posibilidad de utilizar la función denominada *división matricial* \boxdiv , que en forma "automática" resuelve sistemas de ecuaciones de la forma $ax=b$ donde $x=b\boxdiv a$.

Dicho programa puede consultarse en los anexos¹¹ de este trabajo.

3.2 Variantes de la graduación de Whittaker-Henderson.

1. La forma general de diferencias mixtas es:

$$M = \sum_1^n w_x (v_x - u_x)^2 + h_1 \sum_1^{n-1} (\Delta v_x)^2 + h_2 \sum_1^{n-2} (\Delta^2 v_x)^2 + \dots + h_x \sum_1^{n-x} (\Delta^x v_x)^2$$

Nótese que si $h_x > 0$ y $h_j = 0$ para cualquier otra j , obtenemos la fórmula antes discutida:

¹⁰Ver Anexo Tres.

¹¹Ver Anexo Cuatro.

$$M = \sum_1^n w_x (v_x - u_x)^2 + h_x \sum_1^{n-2} (\Delta^2 v_x)^2$$

2. Forma exponencial para S .

Recordemos que en la fórmula de Whittaker-Henderson S es igual a a un polinomio de grado $z-1$.

Sin embargo nuestra creencia previa puede no estar de acuerdo con que la curva graduada se asemeje a un polinomio.

Por ejemplo, supongamos que estamos graduando fuerzas de mortalidad que creemos que siguen la ley de Makeham, entonces nuestra creencia previa es que:

$$t_x = A + Bc^x$$

Y por lo tanto, S deberá definirse en concordancia con esto.

Vemos que:

$$\Delta t_x = B(c-1)c^x \quad y$$

$$\Delta^2 t_x = B(c-1)^2 c^x = r \Delta t_x$$

Donde $r=(c-1)$, entonces $\Delta^2 t_x - r \Delta t_x$ deberá ser igual a cero, tal que podamos definir a S como:

$$S = \sum_1^{n-2} (\Delta^2 v_x - r \Delta v_x)^2$$

Y es con este resultado que llevaremos a cabo la minimización de M .

3. Variación propuesta por Walter B. Lowrie.

Esta variación consiste en la combinación de las formas polinomial y exponencial de S , su forma es la siguiente:

$$M = (1-h_1) \sum_1^n w_x (v_x - u_x)^2 + h_1 \sum_1^n w_x^s (v_x - s_x)^2 + h_2 \sum_1^{n-2} (\Delta^2 v_x - r \Delta^{2-1} v_x)^2.$$

donde s_x es un valor de una tabla apropiadamente seleccionada y w_x^s es la ponderación asociada con el ajuste entre v_x y s_x .

4. Variación propuesta por D. R. Schuette.

Schuette propone reemplazar las medidas de ajuste y de suavización que se encuentran elevadas al cuadrado en la fórmula de Whittaker-Henderson por valores absolutos, justificándose en el hecho de que el elevar al cuadrado dichas medidas solamente es apropiado en el caso de que el error (v.a.) E_x , se distribuya normalmente, tal como se asumió al desarrollar la fórmula de graduación de Whittaker-Henderson. Cuando este no sea el caso será más conveniente utilizar la siguiente variación:

$$M = \sum_1^n w_x |v_x - u_x| + h \sum_1^{n-2} |\Delta^2 v_x|$$

3.3 Determinación de la fórmula básica de la graduación de Bayes.

Proceso formal de la Graduación de Bayes.

Se especifica que se trata del *proceso formal* de graduación de Bayes dado que existen diversas aplicaciones de métodos que tienen una aproximación de carácter bayesiano.

Más que una fórmula, la graduación de Bayes es un procedimiento a seguir según las nociones propuestas por él.

Por lo tanto, la aproximación Bayesiana estándar al problema de graduación puede ser descrita en cuatro pasos fundamentales:

1.- Formular una *distribución de probabilidad anterior (a priori)* para las t_x , que represente la secuencia que deseamos estimar.

2.- Seleccionar el modelo para el experimento, i.e., una expresión para la *distribución de probabilidad condicional* de los datos observados u_x , dada la secuencia t_x .

3.- Usar el teorema de Bayes para resolver la *distribución de probabilidad posterior (a posteriori)* de t_x , dados los datos obtenidos u_x .

4.- Seleccionar los valores graduados v_x de la distribución posterior de acuerdo al objetivo del problema de graduación.

Descripción de cada uno de los pasos del proceso.

1. La distribución anterior para t_x .

Tomemos a los valores reales que deseamos estimar como variables aleatorias que podamos representar por T_x . En notación vectorial usaremos T para representar el vector aleatorio, y t para el vector de realizaciones de T . Dado que T es un vector de variables aleatorias *continuas*¹², lo que estamos buscando establecer es la *función de densidad anterior* $f_T(t)$, la que por supuesto tendrá que ser una función multivariada.

Teóricamente esta distribución anterior deberá reflejar apropiadamente las *creencias iniciales* acerca de T . Esto es definitivamente mucho más fácil de decir que de realizar, dado que no sólo depende de lo que se tenga *a priori*, sino que además la elección debe poseer ciertas propiedades matemáticas que nos faciliten su manejo (como el caso de las distribuciones conjugadas) y no perdemos al obtener un resultado que no podamos identificar. En cuanto a los cálculos debemos cuidar que no se compliquen demasiado como para no poder resolverlas apropiadamente con una calculadora o computadora i.e., que sólo queden como una teoría difícil de aplicar en la práctica.

2. El modelo para el experimento.

Aquí estamos hablando ni más ni menos que de la función de densidad condicional para los datos observados dado t , i.e. $f_{U/T}(u/t)$.

¹²Es continua dado que para poder graduar suponemos continuidad en el fenómeno que se está estimando.

Este paso es menos subjetivo que el anterior, claro que lo más lógico es suponer que la distribución apropiada de la variable aleatoria U_x es binomial con parámetros t_x y n_x (como la visto en el capítulo 2), pero para nuestros propósitos resultará mucho más cómodo asumir alguna distribución de forma *continua*¹³ que una distribución binomial de forma *discreta*.

3. La distribución posterior para t .

Para encontrar la distribución posterior para t , $f_{T/U}(t/u)$, aplicamos el teorema de Bayes de la siguiente forma:

$$f_{T/U}(t/u) = \frac{f_{U/T}(u/t) \cdot f_T(t)}{f_U(u)}$$

Recordando que $f_{T/U}(t/u)$ es multivariada y continua.

Entonces dado que $f_T(t)$ y $f_{U/T}(u/t)$ son funciones ya dadas encontraremos $f_U(u)$ por medio de:

$$f_U(u) = \int_0^1 f_{U/T}(u/t) \cdot f_T(t) dt$$

que no es otra cosa más que una distribución marginal.

Obteniendo esto automáticamente podremos calcular $f_{T/U}(t/u)$, dado que ya contaremos con todos los elementos necesarios para hacerlo.

Cuando realizamos esta operación es necesario multiplicar $f_T(t)$ y $f_{U/T}(u/t)$, entonces lo mejor será que para obtener una expresión *conveniente* nos auxiliemos de las distribuciones conjugadas, mismas que fueron definidas en el capítulo dos.

¹³Dado que un decremento puede ocurrir en cualquier momento.

El escoger las funciones de distribución tal que sean conjugadas no es absolutamente necesario para nuestro trabajo de graduación Bayesiana, pero sí facilita en mucho las cosas.

4. Selección de los valores graduados v_x .

Concluido el paso anterior tendremos la certeza de que la *distribución posterior* obtenida, contiene nuestra creencia inicial sobre T modificada por la información contenida en u (datos observados).

Es lógico entonces que los valores graduados v se seleccionen de la información contenida en $f_{T|U}(t|u)$.

Recordemos que una función de probabilidad cuenta con medidas de tendencia central llamadas media, mediana y moda que nos describen a la distribución de la población en estudio.

Dado que el objetivo de la graduación es tomar como valores graduados aquellos valores de T esperados, o bien siendo más precisos, tomar aquellos valores más representativos o más probables, podemos tomar en cuenta lo siguiente:

• Si tomamos el criterio de valores esperados, el vector de los valores graduados será igual a la *esperanza (o media)* de la distribución posterior:

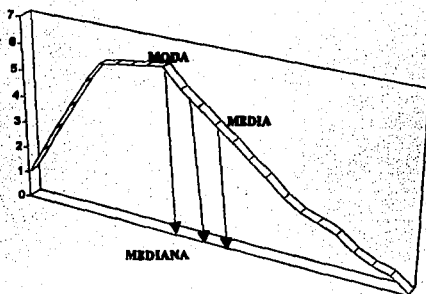
$$E\left[f_{T|U}(t|u)\right]$$

• Si tomamos el criterio de valores representativos (i.e. el valor central de un conjunto de números ordenados por su magnitud), el vector de los valores graduados será igual a la *mediana* de la distribución posterior:

$$F_T(t_{1/2}) = \frac{1}{2}$$

• Por último, si tomamos el criterio de valores más probables (i.e. los valores que ocurren con mayor frecuencia), el vector de los valores graduados será igual a la *moda* de la distribución posterior.

$$\text{máx}\{f_T(t)\}$$



Decidido el criterio a utilizar habremos completado nuestro trabajo de graduación y con esto finalizamos la descripción de cada uno de los pasos del Proceso Formal Bayesiano.

El método de Kimeldorf-Jones.

Este método se auxilia del proceso de graduación formal Bayesiano pero con la característica de que utiliza en sus distribuciones anteriores y posteriores la normal multivariada, por comodidad este será el método que se utilizará para desarrollar el ejemplo práctico en capítulos posteriores. Los pasos se describen a continuación:

1.- La distribución anterior para $f_T(t)$ es:

$f_T(t)$ pertenece a la familia de distribuciones multinormales, lo cual resulta matemáticamente conveniente.

Sean m y A el vector de medias y la matriz de covarianzas respectivamente, entonces la función de densidad para T será de la forma:

$$f_T(t) = k_1 \exp\left[-\frac{1}{2}(t-m)'A^{-1}(t-m)\right]$$

donde:

$$k_1 = [(2\pi)^n |A|]^{-1/2}$$

Esta función se especifica por completo asignándole valores a m y a A , (la elección de A y m es en realidad el eje central de este método).

Es importante notar (debido a la utilización futura de métodos de resolución de sistemas de ecuaciones) que A es simétrica, dado que fue definida como una matriz de covarianzas, de dimensión $n \times n$ y además positiva definida, y por lo tanto no-singular, para que así T resulte multinormal.

2.- El modelo para el experimento.

Como hemos visto la distribución que mejor representa a U_x dado t_x , es la Binomial (o proporción Binomial Condicional). Dado que una proporción binomial se aproxima a una Normal, el método de Kimeldorf-Jones utiliza como distribución condicional multivariada para U , dado t , una multinormal.

Como $E[U_x] = t_x$, se sigue que el vector de medias para nuestra multinormal condicional es t . Si suponemos además independencia mutua entre las variables U_x , las covarianzas serán todas cero y por lo tanto la matriz de covarianzas, B , es una matriz diagonal $n \times n$, que contiene solamente las varianzas de las variables U_x .

Lo cual significa que la función de densidad condicional para el modelo queda expresada por:

$$f_{U|T}(u|t) = k_2 \exp\left[-\frac{1}{2}(u-t)'B^{-1}(u-t)\right]$$

donde:

$$k_2 = [(2\pi)^n |B|]^{-1/2}$$

3.- La distribución posterior para T.

Para obtener esta distribución aplicaremos el teorema de Bayes.

Ya conocemos $f_{U|T}(u|t)$ y $f_T(t)$ entonces $f_U(u)$ lo podemos obtener por medio de:

$$f_U(u) = \int_0^1 f_{U|T}(u|t) \cdot f_T(t) dt$$

Debido a que $f_U(u)$ es una integral definida, podemos concluir que es constante con respecto a t, y por lo tanto podemos combinarla con k_1 y k_2 de la siguiente forma, sean:

$$k_3 = f_U(u) \text{ y } k_4 = \frac{k_1 \cdot k_2}{k_3},$$

ahora sí, aplicando el teorema de Bayes obtenemos:

$$\begin{aligned} f_{T|U}(t|u) &= \frac{k_2 \exp\left[-\frac{1}{2}(u-t)' B^{-1}(u-t)\right] \cdot k_1 \exp\left[-\frac{1}{2}(t-m)' A^{-1}(t-m)\right]}{k_3} \\ &= \frac{k_1 k_2}{k_3} \cdot \exp\left\{-\frac{1}{2}\left[(u-t)' B^{-1}(u-t) + (t-m)' A^{-1}(t-m)\right]\right\} \\ &= k_4 \cdot \exp\left\{-\frac{1}{2}\left[(u-t)' B^{-1}(u-t) + (t-m)' A^{-1}(t-m)\right]\right\} \end{aligned}$$

Y reorganizando este exponente;

$$\begin{aligned}
 f_{T|U}(t|u) &= k_4 \exp\left\{-\frac{1}{2}\left[(t-m)'A^{-1}(t-m) + (u-t)'B^{-1}(u-t)\right]\right\} \\
 &= k_4 \exp\left\{-\frac{1}{2}\left[(t'-m')A^{-1}(t-m) + (u'-t')B^{-1}(u-t)\right]\right\} \\
 &= k_4 \exp\left\{-\frac{1}{2}\left[(t'A^{-1}-m'A^{-1})(t-m) + (u'B^{-1}-t'B^{-1})(u-t)\right]\right\} \\
 &= k_4 \exp\left\{-\frac{1}{2}\left[t'A^{-1}t - t'A^{-1}m - m'A^{-1}t + m'A^{-1}m + u'B^{-1}u \dots \right.\right. \\
 &\quad \left.\left. \dots - u'B^{-1}t - t'B^{-1}u + t'B^{-1}t\right]\right\} \\
 &= k_4 \exp\left\{-\frac{1}{2}\left[(m'A^{-1}m + u'B^{-1}u) + t'A^{-1}t - t'A^{-1}m - m'A^{-1}t \dots \right.\right. \\
 &\quad \left.\left. \dots - u'B^{-1}t - t'B^{-1}u + t'B^{-1}t\right]\right\} \\
 &= k_4 \exp\left[-\frac{1}{2}(m'A^{-1}m + u'B^{-1}u)\right] \dots \\
 &\quad \dots \exp\left[-\frac{1}{2}(t'A^{-1}t - t'A^{-1}m - m'A^{-1}t - u'B^{-1}t - t'B^{-1}u + t'B^{-1}t)\right]
 \end{aligned}$$

Sea entonces:

$$k_5 = k_4 \exp\left[-\frac{1}{2}(m'A^{-1}m + u'B^{-1}u)\right]$$

Sustituyendo y reordenando los términos restantes obtenemos:

$$k_5 \cdot \exp\left\{-\frac{1}{2}\left[t'(A^{-1} + B^{-1})t - t'(B^{-1}u + A^{-1}m) - (u'B^{-1} + m'A^{-1})t\right]\right\}$$

Ahora bien definamos a

$$C^{-1} = (A^{-1} + B^{-1})$$

y a

$$v = C(B^{-1}u + A^{-1}m),$$

tal que

$$C^{-1}v = (B^{-1}u + A^{-1}m),$$

y sustituyendo esto en nuestro resultado anterior obtenemos:

$$k_5 \cdot \exp\left\{-\frac{1}{2}\left[t'C^{-1}t - t'C^{-1}v - (u'B^{-1} + m'A^{-1})t\right]\right\} \dots @$$

Dado que $C^{-1}v = (B^{-1}u + A^{-1}m)$, entonces:

$$(C^{-1}v)' = v'(C^{-1})' = (B^{-1}u)' + (A^{-1}m)' = u'(B^{-1})' + m'(A^{-1})'$$

Observemos que para cualquier matriz X diagonal se cumple:

$$(X^{-1})' = X^{-1},$$

ya que las matrices A , B y C son diagonales por definición, tenemos que:

$$v'C^{-1} = u'B^{-1} + m'A^{-1}$$

Entonces sustituyendo $v'C^{-1}$ por $u'B^{-1} + m'A^{-1}$ en ...@ y sumando y restando $v'C^{-1}v$ dentro del paréntesis obtenemos:

$$\begin{aligned} & k_5 \cdot \exp\left\{-\frac{1}{2}\left[t'C^{-1}t - t'C^{-1}v - v'C^{-1}t\right]\right\} \\ &= k_5 \cdot \exp\left\{-\frac{1}{2}\left[t'C^{-1}t - t'C^{-1}v - v'C^{-1}t + v'C^{-1}v - v'C^{-1}v\right]\right\} \end{aligned}$$

Sea $k_6 = k_5 \cdot \exp\left\{-\frac{1}{2}v'C^{-1}v\right\}$ y notemos además que:

$$(t-v)'C^{-1}(t-v) = t'C^{-1}t - t'C^{-1}v - v'C^{-1}t + v'C^{-1}v$$

$$\therefore f_{TV}(t|u) = k_6 \cdot \exp\left\{-\frac{1}{2}(t-v)'C^{-1}(t-v)\right\}$$

donde:

1. k_6 no involucra a t .

$$2. \quad v = (A^{-1} + B^{-1})^{-1}(B^{-1}u + A^{-1}m), \text{ y}$$

$$3. \quad C = (A^{-1} + B^{-1})^{-1}.$$

Con esto demostramos que $f_{TV}(t|u)$ es también una distribución multinormal con vector de medias v y matriz de covarianzas C , es decir una multinormal es conjugada consigo misma, lo cual simplifica nuestro trabajo.

4.- La selección de los valores graduados.

Como ya habíamos dicho lo ideal es escoger los valores graduados como aquellos representados por la media, moda o mediana de la distribución posterior $f_{TV}(t|u)$, el estar utilizando el método de Kimeldorf-Jones, con distribuciones multinormales, nos conlleva a que estas tres medidas coinciden en un mismo punto, siendo v el vector de medias, modas y medianas, tenemos que nuestros valores graduados son:

$$\begin{aligned} v &= (A^{-1} + B^{-1})^{-1}(B^{-1}u + A^{-1}m) \\ &= u + (I + AB^{-1})^{-1}(m - u), \end{aligned}$$

donde I es una matriz identidad de orden n .

Ahora bien para encontrar estos valores graduados necesitamos especificar primeramente los elementos que integran a m , A y B . A continuación se presentan algunas sugerencias para su elección:

- m representa la media, moda y mediana de T cuando esta se distribuye normalmente, pero también es nuestra mejor estimación de t en ausencia de datos experimentales.

Entonces, si no contamos con ninguna información *anterior*, lo que podemos hacer es seleccionar estimaciones ya existentes que estén basadas en la experiencia de otro grupo (muestra) con características similares al grupo que tenemos en estudio actualmente.

• A es la matriz de covarianzas del vector aleatorio T . Recordando que se definió a la teoría de la graduación como las "relaciones existentes entre tasas vecinas", podemos representar estas relaciones por medio de coeficientes¹⁴ de correlación entre cada par de tasas.

Claramente los elementos de la diagonal de A son las varianzas de las variables aleatorias de T , y los demás deberán ser las covarianzas. Como definitivamente la elección de A es bastante subjetiva nos resultará más útil el concepto de coeficiente de correlación que el de covarianza. Así, pues, determinamos que este coeficiente siempre será mayor o igual que cero y que tenderá a cero en tanto que la distancia entre los elementos de T sea mayor.

De acuerdo a esto podremos igualar a cero ciertos elementos de A con el propósito de simplificar los cálculos.

En el presente trabajo, para no sobrecargar la teoría, tomaremos como tal la fórmula sugerida por Kimeldorf y Jones que nos dice:

$$a_{ij} = p^2 r^{|i-j|}, \quad p > 0, \quad 0 \leq r < 1$$

donde:

- 1.- $a_{ij} \in A$
- 2.- p^2 es la varianza de $T_i \quad \forall i$.
- 3.- r es el coeficiente de correlación entre T_i y T_j donde $|i-j|=1$.
- 4.- De igual forma, si $|i-j|=2$, el coeficiente de correlación entre T_i y T_j será r^2 , etc.

¹⁴Coeficiente de correlación.- Es el grado de interconexión entre variables que intenta determinar con que precisión describe o explica la relación entre variables una ecuación lineal o de cualquier otro tipo.

Observemos que:

- 1.- A será definida positiva y claramente simétrica.
- 2.- Asumimos una varianza igual (p^2) para toda T_i .
- 3.- La correlación entre los elementos de A se establece en forma geométrica ($r^{|i-j|}$).

Esto no significa que debamos limitar nuestra creatividad si no estamos de acuerdo a esta fórmula propuesta. En el caso de esta tesis se adoptará, como ya dijimos, tal y como está para no sobrecargar la teoría y no abusar en el tiempo de realización de la misma.

Pero bien se puede suponer varianzas diferentes para cada T_i al observar que nuestra muestra presente mayor sesgo en ciertas edades, hablando de la construcción de una tabla de mortalidad graduada, o dificultad para obtener información adecuada en ciertos rangos.

O en general, construir una matriz de covarianzas más acorde a nuestras creencias o necesidades¹⁵.

Pero sobre todo estar siempre consciente de que no existe la graduación *correcta* en términos absolutos, de lo contrario nos encontraremos vagando en un gran callejón sin salida.

• B es una matriz diagonal de covarianzas, o mejor dicho varianzas, de la distribución $f_{U|T}(u|t)$, sus elementos son más fáciles de seleccionar en comparación con los de A .

El que tengamos la creencia (lógica) de que U_i es aproximadamente una proporción binomial nos permite calcular:

¹⁵Por ejemplo en las aseguradoras para evitar pérdidas financieras substanciales sería preferible utilizar tablas de mortalidad *sobrevaluadas* (mayor mortalidad, previendo mayor cantidad de pagos reclamados). Sin embargo en la creación de un plan de pensiones por jubilación sería preferible optar por una tabla *subvaluada* (mayor esperanza de vida, previendo un periodo más largo durante el que se tendrán que hacer los pagos).

$$\text{Var}(U_i) = \frac{t_i(1-t_i)}{n_i}$$

Pero como en el método de Kimeldorf-Jones B debe estar libre de t , debemos de tomar:

$$\text{Var}(U_i) = b_{ii} = \frac{m_i(1-m_i)}{n_i}$$

dado que m es nuestra mejor estimación disponible de T .

Contando con estos 3 parámetros, habremos por fin resuelto nuestro problema de graduación por medio de este método. En los anexos se presenta el programa en APL para llevar a cabo el ejemplo práctico presentado en este trabajo.

3.4 Variantes de la graduación de Bayes.

Para este caso, está de más decir que existen tantas variantes como nuestra habilidad nos permita crear.

Todo dependerá de las distribuciones escogidas *a priori* y *posteriori* para obtener los valores graduados.

CAPITULO CUATRO.

DIFERENCIAS ANALITICAS ENTRE LAS DOS TECNICAS.

4.1 Diferencias en la concepción

La técnica de graduación de Whittaker-Henderson fue concebida como una modificación a las técnicas de regresión lineal, como ya se vió. Solamente toma en cuenta las actuales observaciones del fenómeno en estudio.

Por otro lado la técnica de graduación de Bayes nace de la definición de probabilidad condicional; i.e. se interesa por los hechos históricos para modificar los eventos actuales.

4.2 Diferencias en las herramientas utilizadas.

La técnica de graduación de Whittaker-Henderson se auxilia de la estadística, de conceptos como varianza, de los procedimientos de regresión lineal y de conceptos de métodos numéricos (diferencias finitas).

Mientras que la graduación de Bayes es una técnica netamente de la Estadística Bayesiana. La cual implica un amplio conocimiento de las distribuciones de probabilidad para poder sacarle ventaja a una técnica tan creativa como lo puede ser esta.

Por lo tanto las herramientas matemáticas utilizadas en la graduación de Whittaker-Henderson suelen ser más sencillas que aquellas utilizadas en la graduación de Bayes que implica herramientas de cálculo más complejas.

4.3 Ventajas y desventajas, similitudes y diferencias generales.

La graduación de Whittaker-Henderson tiene una gran ventaja práctica:

Su estructura es bastante accesible en cuanto a cálculos de resultados, i.e., el hecho de que finalmente se llegue a la solución de un sistema de ecuaciones nos pinta un panorama por demás conocido.

La única dificultad se tendría al aplicar el algoritmo apropiado del álgebra lineal para resolver dicho sistema de ecuaciones.

La desventaja consiste en que como sólo se basa en observaciones actuales, no es posible aplicar correcciones que nos permitan uniformar nuestros valores con aquellos ya existentes. Por otro lado nuestra creatividad se encuentra limitada por su misma estructura.

La graduación de Bayes tiene una gran ventaja sobre la de Whittaker-Henderson, el poder incluir en sus valores graduados la información de valores ya existentes que nos uniformen dichas estimaciones.

Sin embargo en la práctica resulta muy densa de realizar debido a su naturaleza teórica y el escoger los parámetros y distribuciones apropiados que nos lleven a un buen resultado. Por otro lado nos permite experimentar con una gran variedad de posibilidades al graduar por medio de las llamadas distribuciones conjugadas.

En pocas palabras la graduación de Bayes incluye un grado de subjetividad mayor que aquel que aplicamos para la graduación de Whittaker-Henderson.

Definitivamente la similitud más relevante entre estas dos técnicas es su objetivo:

Graduar valores presentados en forma tabular que no incluyen esa relación entre vecindades de los datos que nosotros intuimos existe.

CAPITULO CINCO. APLICACIONES NUMERICAS.

5.1 Tablas utilizadas.

Lo primero que debemos obtener para poder graduar es una tabla de tasas brutas.

Se estuvo buscando un trabajo de esta naturaleza que se hubiera realizado para el año de 1990 en base al censo de población y número de defunciones en dicho año.

Al no haberse encontrado una tabla de mortalidad hecha para este año, para nuestro país, en base a este censo, se procedió a calcular las tasas brutas de manera muy simple i.e., sólo para efectos de ilustración del presente trabajo, ya que no es su objetivo principal el de construir dichas tablas.

Esto se llevó a cabo de la siguiente forma:

1. Intervinieron en el cálculo datos provenientes de:
 - a. El censo de población mexicano de 1990.
 - b. Defunciones en nuestro país *registradas* en 1990.

2. Debido a que las defunciones originalmente venían agrupadas por quinquenio se procedió a interpolar linealmente tomando como referencia las marcas de clase de cada intervalo.

Ejemplo:

<i>Intervalo de edad</i>	<i>Marca de clase</i>	<i>Defunciones</i>
40-44	42	8776
45-49	47	10204

$$m = \frac{8776 - 10204}{42 - 47} = 285.60$$

$$y - 8776 = 285.60(x - 42)$$

$$y = 285.60(x - 42) + 8776 \dots(1)$$

Por lo tanto el número de muertes a edades 43, 44, 45 y 46 se obtendrán por sustitución en la ecuación (1), e.g.;

A edad 44 tendremos:

$$y = 285.60(44 - 42) + 8776$$

$$y = 9347$$

Como no es posible tener 1/5 de muerto, los resultados fueron redondeados a números enteros.

3. Las tasas brutas se calcularon dividiendo el número de defunciones a edad x entre el número de vivos a edad x en 1990, al observar los resultados se observó que aquellas tasas obtenidas a partir de edad 86 resultaban mayores a uno, por lo cual fueron desechadas.

4. También se utilizó para el caso de la graduación de Bayes la tabla de mortalidad AMA91 elaborada por el departamento de Estadística y Actuaría del ITAM en base a la información proporcionada por 5 de las aseguradoras más importantes en nuestro país.

5. Las tabulaciones de toda esta información y sus respectivas gráficas son mostradas en las páginas siguientes.

6. Por último se remarca que se hizo una distinción entre mujeres y hombres para la presentación de las ejemplificaciones.

TABLA 1

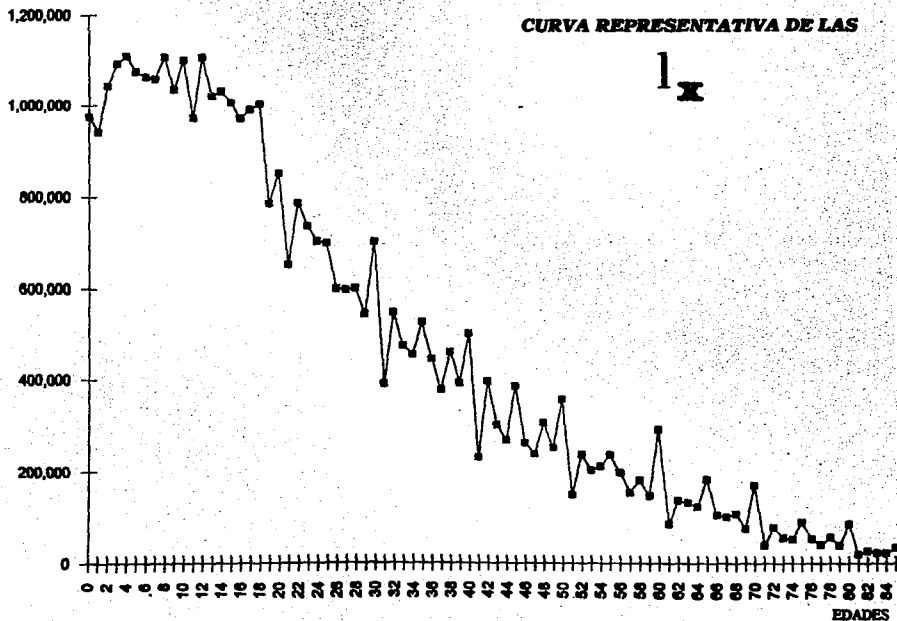
HOMBRES

EDADES	l_x	d_x	q_x	AMAS1 q_x
0	975,043	38,768	0.037707	0.008300
1	940,971	5,397	0.005736	0.000339
2	1,043,565	2,539	0.002433	0.000447
3	1,091,225	1,570	0.001439	0.000436
4	1,109,198	1,116	0.001008	0.000435
5	1,074,622	1,943	0.001808	0.000439
6	1,062,813	2,771	0.002607	0.000442
7	1,059,395	3,598	0.003396	0.000447
8	1,105,859	3,532	0.003194	0.000434
9	1,035,596	3,468	0.003347	0.000418
10	1,099,429	3,400	0.003093	0.000390
11	972,013	3,334	0.003430	0.000380
12	1,105,953	3,268	0.002955	0.000308
13	1,020,991	3,950	0.003869	0.000439
14	1,032,272	4,832	0.004487	0.000587
15	1,008,138	5,313	0.005281	0.000701
16	972,458	5,995	0.006185	0.000848
17	990,507	6,677	0.006741	0.001015
18	1,003,619	7,057	0.007032	0.001023
19	787,172	7,438	0.009448	0.001051
20	852,582	7,816	0.009187	0.001093
21	653,448	8,195	0.012541	0.001140
22	787,887	8,575	0.010882	0.001185
23	738,809	8,924	0.011673	0.001257
24	705,302	8,872	0.012295	0.001321
25	702,150	8,721	0.012420	0.001378
26	602,037	8,789	0.014566	0.001431
27	568,739	8,818	0.014728	0.001480
28	603,185	8,758	0.014518	0.001493
29	544,484	8,894	0.015967	0.001502
30	705,371	8,833	0.012239	0.001507
31	392,429	8,571	0.021841	0.001505
32	548,253	8,509	0.015520	0.001498
33	477,121	8,807	0.018039	0.001573
34	455,562	8,705	0.019108	0.001645
35	528,480	8,802	0.016655	0.001718
36	448,194	8,900	0.019948	0.001789
37	379,555	8,998	0.023707	0.001888
38	482,145	8,954	0.019375	0.002023
39	394,191	8,909	0.022601	0.002190
40	503,551	8,865	0.017805	0.002373
41	231,940	8,820	0.038027	0.002573
42	397,684	8,778	0.022068	0.002794
43	301,465	9,082	0.030060	0.003018
44	270,373	9,347	0.034571	0.003284
45	387,699	9,833	0.024847	0.003532
46	263,689	9,918	0.037584	0.003821

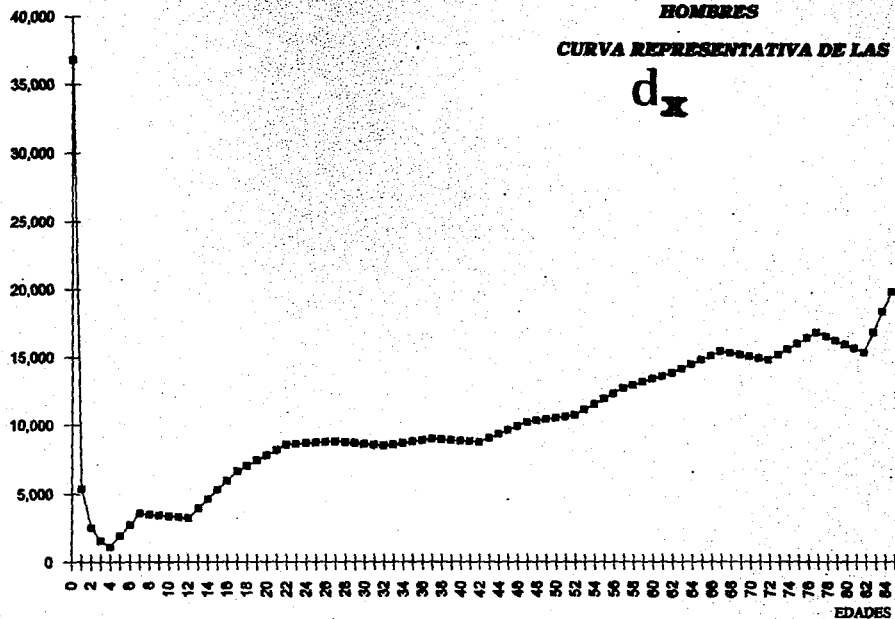
TABLE 1

47	240,341	10,204	0.042456	0.004130
48	306,568	10,311	0.033634	0.004470
49	254,076	10,417	0.041000	0.004831
50	357,576	10,524	0.029432	0.005211
51	149,507	10,630	0.071100	0.005613
52	238,486	10,737	0.045021	0.006036
53	203,630	11,133	0.054673	0.006638
54	212,674	11,530	0.054214	0.007269
55	236,690	11,628	0.050344	0.007934
56	199,115	12,323	0.061889	0.008639
57	154,091	12,719	0.082542	0.009362
58	181,607	12,941	0.071258	0.010542
59	147,161	13,163	0.089446	0.011760
60	263,316	13,366	0.045637	0.013058
61	85,090	13,608	0.159625	0.014451
62	136,396	13,830	0.101396	0.015951
63	132,174	14,152	0.107071	0.017957
64	122,941	14,475	0.117739	0.020100
65	162,784	14,797	0.080953	0.022395
66	103,911	15,120	0.145509	0.024859
67	100,024	15,442	0.154383	0.027505
68	106,204	15,313	0.144185	0.030514
69	74,718	15,184	0.203217	0.033727
70	170,276	15,054	0.088409	0.037153
71	36,434	14,925	0.378480	0.040796
72	76,970	14,796	0.192231	0.044664
73	55,009	15,203	0.276373	0.050001
74	52,342	15,010	0.286231	0.055619
75	88,881	16,018	0.180216	0.061566
76	52,993	16,425	0.309947	0.067893
77	40,679	16,632	0.411752	0.074648
78	56,980	16,543	0.295516	0.079757
79	36,102	16,259	0.415656	0.085260
80	84,976	15,964	0.187865	0.091072
81	20,243	15,674	0.774262	0.097110
82	27,802	15,385	0.553377	0.103269
83	23,392	16,663	0.720667	0.110249
84	23,407	16,342	0.763612	0.117295
85	34,937	16,820	0.567307	0.124455

NUMERO DE VIVOS

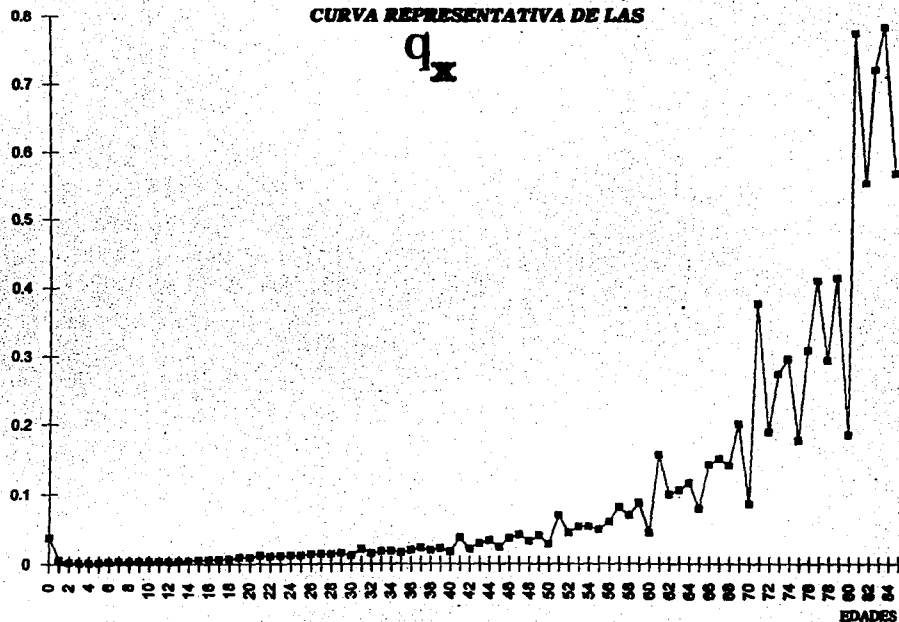


NUMERO DE DEFUNCIONES



PROB. DE MUERTE

HOMBRES
CURVA REPRESENTATIVA DE LAS
 q_x



PROB. DE MUERTE

HOMBRES

AMA 91

q_x

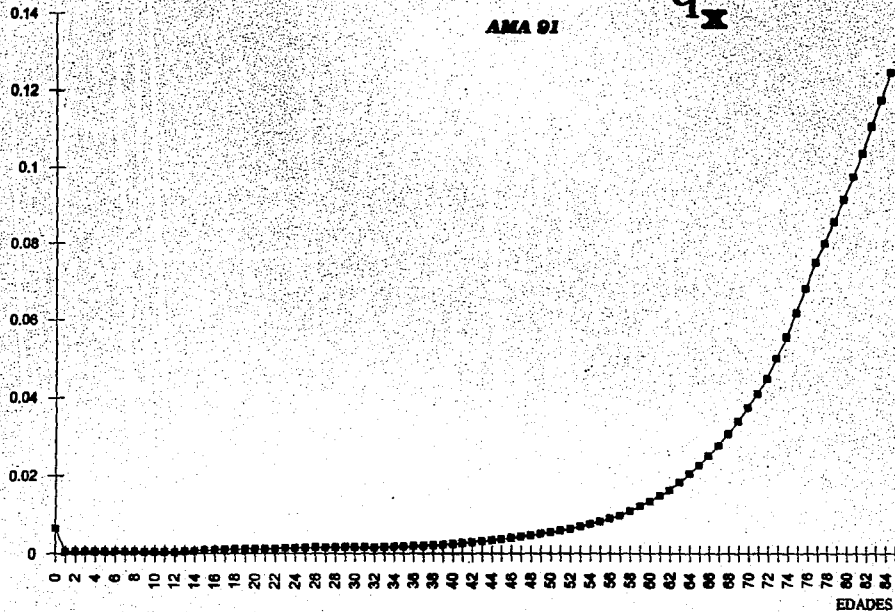


TABLA 2

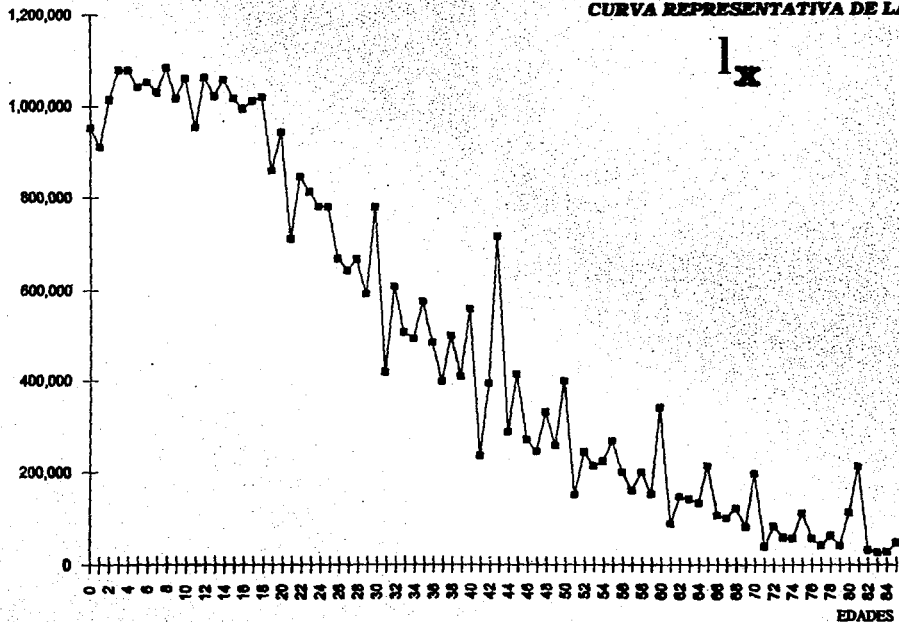
MUJERES

EDADES	l_x	d_x	q_x	ANAS1 q_x
0	852,784	28,396	0.028506	0.004200
1	811,777	4,809	0.005274	0.000327
2	1,014,109	2,313	0.002261	0.000418
3	1,078,513	1,395	0.001293	0.000335
4	1,077,993	948	0.000878	0.000381
5	1,041,326	1,585	0.001522	0.000371
6	1,052,355	2,225	0.002114	0.000359
7	1,030,111	2,854	0.002771	0.000339
8	1,083,608	2,720	0.002510	0.000327
9	1,018,549	2,575	0.002533	0.000303
10	1,060,871	2,431	0.002292	0.000287
11	954,008	2,296	0.002396	0.000221
12	1,063,337	2,142	0.002014	0.000185
13	1,022,215	2,291	0.002241	0.000249
14	1,058,203	2,440	0.002306	0.000330
15	1,017,598	2,590	0.002545	0.000414
16	995,104	2,739	0.002752	0.000507
17	1,012,127	2,888	0.002853	0.000614
18	1,020,436	2,932	0.002873	0.000593
19	859,248	2,977	0.003485	0.000587
20	843,181	3,021	0.003203	0.000590
21	710,231	3,066	0.004317	0.000595
22	844,950	3,110	0.003681	0.000596
23	811,831	3,125	0.003849	0.000641
24	780,842	3,140	0.004021	0.000677
25	779,870	3,154	0.004045	0.000708
26	669,409	3,189	0.004734	0.000736
27	644,261	3,184	0.004942	0.000763
28	668,233	3,217	0.004814	0.000810
29	592,344	3,249	0.005485	0.000858
30	779,884	3,282	0.004208	0.000909
31	419,800	3,314	0.007894	0.000963
32	608,199	3,347	0.005503	0.001021
33	507,593	3,513	0.006921	0.001115
34	493,407	3,678	0.007454	0.001216
35	573,950	3,844	0.006697	0.001323
36	484,691	4,009	0.008271	0.001440
37	399,857	4,175	0.010441	0.001566
38	499,488	4,237	0.008483	0.001692
39	410,565	4,299	0.010471	0.001830
40	557,258	4,362	0.007828	0.001978
41	235,990	4,424	0.018747	0.002135
42	394,795	4,886	0.012376	0.002302
43	717,705	4,694	0.006540	0.002445
44	287,011	4,902	0.017079	0.002595
45	415,071	5,111	0.012314	0.002752
46	270,502	5,319	0.019863	0.002914

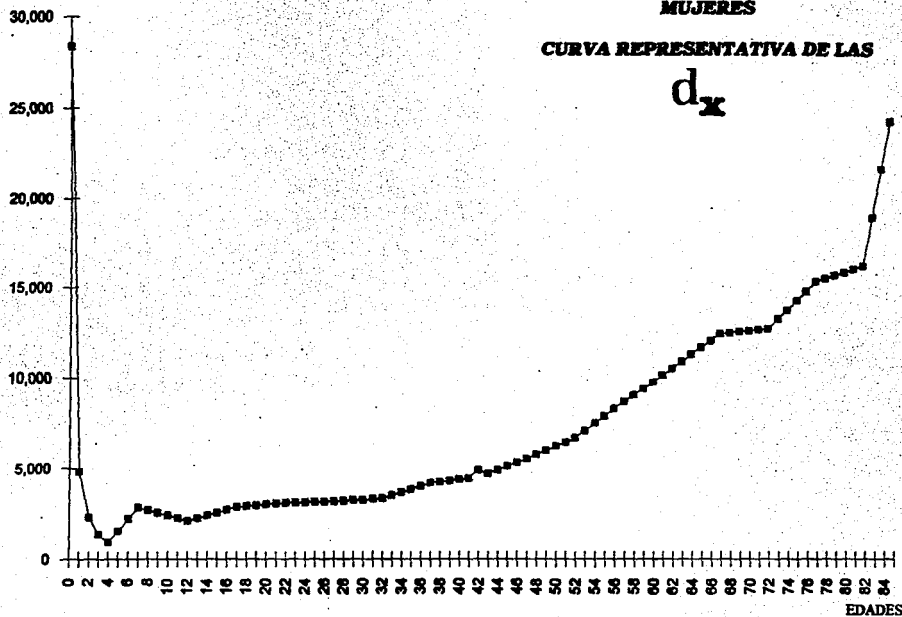
TABLE 2

47	245,794	5,527	0.022466	0.003080
48	329,741	5,762	0.017474	0.003430
49	258,179	5,996	0.023232	0.003791
50	399,290	6,233	0.015610	0.004166
51	149,779	6,469	0.043190	0.004570
52	244,471	6,704	0.027422	0.005003
53	214,200	7,102	0.033156	0.005360
54	224,176	7,500	0.033456	0.005752
55	267,284	7,897	0.029545	0.006174
56	199,858	8,295	0.041504	0.006620
57	169,747	8,693	0.054760	0.007086
58	199,318	9,064	0.048475	0.007941
59	150,413	9,436	0.062734	0.008826
60	339,365	9,807	0.028696	0.009756
61	84,984	10,179	0.119775	0.010746
62	144,654	10,550	0.072933	0.011812
63	140,806	10,920	0.077964	0.012324
64	131,761	11,290	0.065685	0.012901
65	214,043	11,661	0.054480	0.013516
66	103,761	12,031	0.115949	0.014145
67	96,933	12,401	0.126347	0.014762
68	119,831	12,454	0.103930	0.016364
69	79,442	12,507	0.157436	0.017966
70	196,500	12,560	0.063919	0.019617
71	39,644	12,613	0.316560	0.021349
72	80,746	12,666	0.156662	0.023205
73	56,667	13,194	0.224820	0.027033
74	57,219	13,722	0.239615	0.031069
75	109,133	14,250	0.130575	0.035473
76	57,223	14,778	0.256253	0.040226
77	42,214	15,306	0.362581	0.045437
78	62,149	15,474	0.246982	0.049749
79	42,282	15,641	0.369921	0.054530
80	111,104	15,809	0.142260	0.058725
81	212,200	15,976	0.075267	0.063277
82	33,544	16,144	0.481278	0.071130
83	27,623	16,845	0.682221	0.076048
84	26,521	21,546	0.755443	0.085243
85	46,274	24,248	0.502269	0.092747

NUMERO DE VIVOS



NÚMERO DE DEFUNCIONES

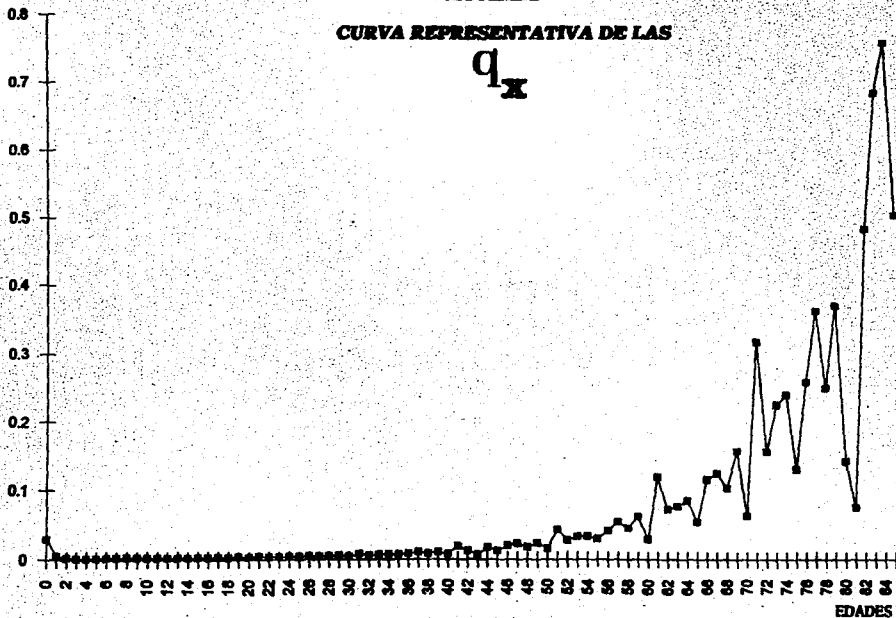


PROB. DE MUERTE

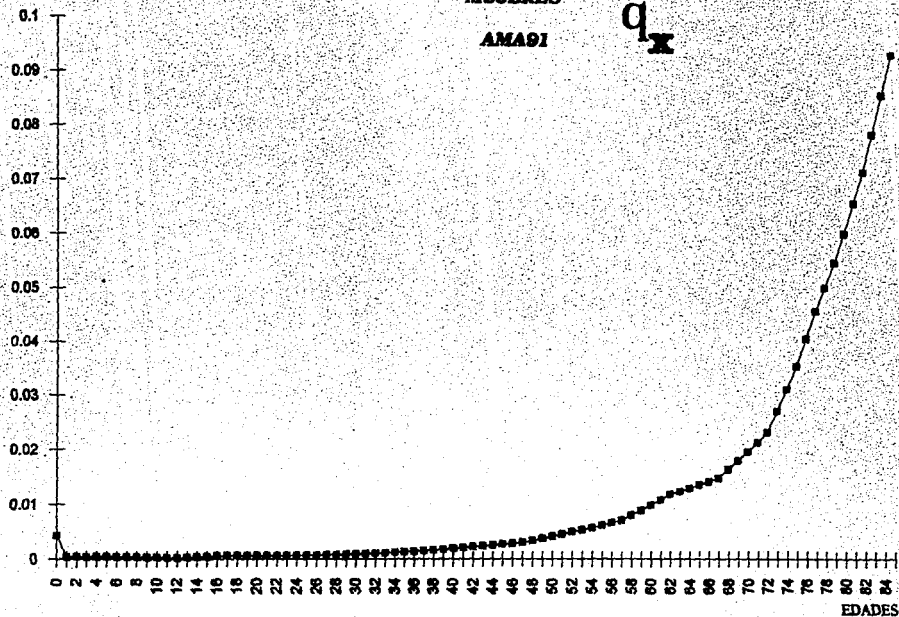
MUJERES

CURVA REPRESENTATIVA DE LAS

q_x



PROB. DE MUERTE



DEFUNCIONES REGISTRADAS POR GRUPO DE EDAD

GRUPOS DE EDAD	NUMERO DE DEFUNCIONES	
	HOMBRES	MUJERES
0	36,766	28,399
1	5,397	4,809
2	2,539	2,313
3	1,570	1,395
4	1,116	946
5-9	3,598	2,864
10-14	3,268	2,142
15-19	6,677	2,888
20-24	8,575	3,110
25-29	8,818	3,184
30-34	8,509	3,347
35-39	8,998	4,175
40-44	8,776	4,486
45-49	10,204	5,527
50-54	10,737	6,704
55-59	12,719	8,693
60-64	13,830	10,550
65-69	15,442	12,401
70-74	14,796	12,666
75-79	16,832	15,306
80-84	15,385	16,144
85 Y MAS	22,777	29,650

5.2 Graduación de las tablas por medio de las dos técnicas.

A continuación se muestran los resultados de las graduaciones que se practicaron.

Dichos resultados fueron posibles gracias a los programas en APL que se muestran en los anexos.

Parámetros utilizados.

Para la graduación de Whittaker-Henderson:

$$z=30 \text{ y } h=0.5$$

Para la graduación de Bayes:

$$b_{ij} = \frac{m_x(1-m_x)}{n_x}; r=0.9 \text{ limitado a la vecindad entre 10 edades adyacentes}$$

$p = \sum \frac{b_{ij}}{n}$; donde m es el vector de probabilidades de muerte de la tabla AMA91, r el coeficiente de correlación, n el número total de observaciones.

HOMBRES

EDADES	TOTAL	MUERTES	T.BRUTAS	GRAD. W-H	GRAD. BAYES	AMA91
0	675,043	36,786	0.037707	0.037707	0.022004	0.006300
1	940,971	5,397	0.005736	0.005736	0.003037	0.000339
2	1,043,665	2,539	0.002433	0.002433	0.001440	0.000447
3	1,091,225	1,570	0.001439	0.001439	0.000937	0.000436
4	1,109,198	1,116	0.001008	0.001008	0.000721	0.000435
5	1,074,822	1,943	0.001808	0.001807	0.001124	0.000439
6	1,062,813	2,771	0.002607	0.002613	0.001525	0.000442
7	1,059,395	3,598	0.003396	0.003373	0.001922	0.000447
8	1,105,859	3,532	0.003194	0.003247	0.001814	0.000434
9	1,035,596	3,466	0.003347	0.003225	0.001882	0.000418
10	1,099,429	3,400	0.003093	0.003270	0.001741	0.000390
11	972,013	3,334	0.003430	0.003135	0.001895	0.000380
12	1,105,953	3,268	0.002955	0.003176	0.001631	0.000306
13	1,020,991	3,950	0.003869	0.003672	0.002154	0.000439
14	1,032,272	4,632	0.004487	0.004515	0.002527	0.000567
15	1,006,136	5,313	0.005281	0.005399	0.002991	0.000701
16	972,458	5,995	0.006185	0.006072	0.003506	0.000848
17	990,507	6,677	0.006741	0.006611	0.003878	0.001015
18	1,003,819	7,057	0.007032	0.007389	0.004027	0.001023
19	787,172	7,436	0.009446	0.008614	0.005249	0.001051
20	852,582	7,816	0.009167	0.010017	0.005130	0.001093
21	853,448	8,195	0.012541	0.011086	0.006841	0.001140
22	787,987	8,575	0.010882	0.011544	0.008034	0.001185
23	738,809	8,624	0.011673	0.011611	0.008485	0.001257
24	705,302	8,672	0.012295	0.011882	0.008808	0.001321
25	702,150	8,721	0.012420	0.012826	0.008899	0.001376
26	602,037	8,769	0.014586	0.014186	0.007966	0.001431
27	596,739	8,818	0.014728	0.015036	0.008104	0.001480
28	603,185	8,756	0.014516	0.014776	0.008005	0.001493
29	544,484	8,694	0.015967	0.014052	0.008735	0.001502
30	705,371	8,633	0.012239	0.014210	0.008673	0.001507
31	382,429	8,571	0.021841	0.015675	0.011673	0.001505
32	548,253	8,509	0.015520	0.017323	0.008508	0.001496
33	477,121	8,507	0.018039	0.017688	0.009806	0.001573
34	455,562	8,705	0.019108	0.017572	0.010377	0.001645
35	528,480	8,802	0.016655	0.017949	0.009186	0.001716
36	448,194	8,900	0.019948	0.018680	0.010888	0.001789
37	379,555	8,998	0.023707	0.021439	0.012787	0.001868
38	482,145	8,954	0.019375	0.021439	0.010899	0.002023
39	394,181	8,909	0.022601	0.020029	0.012385	0.002190
40	503,551	8,865	0.017605	0.019842	0.009969	0.002373
41	231,840	8,820	0.038027	0.022657	0.020300	0.002573
42	397,984	8,778	0.022066	0.026700	0.012431	0.002794
43	301,465	9,062	0.030060	0.028553	0.018957	0.003018
44	270,373	9,347	0.034671	0.027995	0.021715	0.003284

TABLE 4

45	387,699	9,633	0.024847	0.028955	0.016097	0.003532
46	263,689	9,918	0.037584	0.033897	0.023718	0.003821
47	240,341	10,204	0.042456	0.038966	0.026715	0.004130
48	306,568	10,311	0.033634	0.036367	0.021657	0.004470
49	254,076	10,417	0.041000	0.033619	0.026145	0.004831
50	357,578	10,524	0.029432	0.033711	0.019485	0.005211
51	149,507	10,630	0.071100	0.043048	0.044200	0.005613
52	238,488	10,737	0.045021	0.053470	0.029008	0.006036
53	203,630	11,133	0.054673	0.054321	0.034942	0.006638
54	212,674	11,530	0.054214	0.046660	0.034930	0.007269
55	236,690	11,926	0.050344	0.051002	0.032923	0.007934
56	199,115	12,323	0.061889	0.066320	0.040015	0.008639
57	154,091	12,719	0.082542	0.079772	0.052493	0.009392
58	181,607	12,941	0.071258	0.075310	0.046317	0.010542
59	147,161	13,163	0.089446	0.059777	0.057534	0.011760
60	293,318	13,366	0.045637	0.057425	0.032254	0.013058
61	85,090	13,608	0.159925	0.079116	0.100166	0.014451
62	136,396	13,830	0.101396	0.106768	0.066296	0.015951
63	132,174	14,152	0.107071	0.114752	0.070464	0.017957
64	122,941	14,475	0.117739	0.101535	0.077630	0.020100
65	182,784	14,797	0.080953	0.093431	0.056698	0.022395
66	103,911	15,120	0.145509	0.114578	0.095947	0.024859
67	100,024	15,442	0.154383	0.154296	0.102263	0.027505
68	106,204	15,313	0.144185	0.170050	0.097490	0.030514
69	74,718	15,184	0.203217	0.138377	0.133593	0.033727
70	170,276	15,054	0.088409	0.105124	0.067354	0.037153
71	39,434	14,925	0.376480	0.145882	0.239763	0.040796
72	76,970	14,796	0.192231	0.244961	0.131612	0.044664
73	55,009	15,203	0.276373	0.280617	0.183382	0.050001
74	52,342	15,610	0.298231	0.216271	0.198569	0.055619
75	66,681	16,018	0.180218	0.207724	0.131477	0.061566
76	52,993	16,425	0.309947	0.307352	0.210514	0.067893
77	40,879	16,832	0.411752	0.347061	0.273273	0.074648
78	55,980	16,543	0.295516	0.341078	0.206885	0.079757
79	39,102	16,253	0.415656	0.366598	0.279933	0.085260
80	64,976	15,964	0.187665	0.194156	0.148103	0.091072
81	20,243	15,674	0.774262	0.764666	0.496113	0.097110
82	27,602	15,365	0.553377	0.555626	0.368486	0.103289
83	23,392	16,663	0.720867	0.720551	0.470044	0.110249
84	23,407	16,342	0.783612	0.783639	0.506696	0.117295
85	34,937	19,820	0.567307	0.567306	0.385388	0.124455

MUJERES

EADES	TOTAL	MUERTES	T.BRUTAS	GRAD. W-H	GRAD. BAYES	AMA91
0	652,784	26,369	0.029806	0.029806	0.017003	0.004200
1	811,777	4,809	0.005274	0.005274	0.002801	0.000327
2	1,014,109	2,313	0.002281	0.002281	0.001349	0.000418
3	1,078,513	1,395	0.001293	0.001293	0.000814	0.000335
4	1,077,993	948	0.000878	0.000878	0.000828	0.000381
5	1,041,326	1,585	0.001522	0.001521	0.000947	0.000371
6	1,052,355	2,225	0.002114	0.002119	0.001237	0.000359
7	1,030,111	2,854	0.002771	0.002749	0.001555	0.000339
8	1,083,608	2,720	0.002510	0.002555	0.001419	0.000327
9	1,016,549	2,575	0.002533	0.002439	0.001418	0.000303
10	1,060,671	2,431	0.002292	0.002415	0.001279	0.000267
11	954,008	2,288	0.002396	0.002231	0.001309	0.000221
12	1,063,337	2,142	0.002014	0.002119	0.001090	0.000185
13	1,022,215	2,291	0.002241	0.002175	0.001245	0.000249
14	1,058,203	2,440	0.002306	0.002323	0.001318	0.000330
15	1,017,596	2,590	0.002545	0.002546	0.001480	0.000414
16	995,104	2,739	0.002752	0.002753	0.001630	0.000507
17	1,012,127	2,888	0.002853	0.002843	0.001734	0.000614
18	1,020,436	2,932	0.002873	0.002937	0.001733	0.000593
19	859,248	2,977	0.003465	0.003202	0.002028	0.000587
20	843,181	3,021	0.003203	0.003549	0.001898	0.000590
21	710,231	3,066	0.004317	0.003768	0.002456	0.000585
22	844,950	3,110	0.003681	0.003842	0.002139	0.000588
23	811,831	3,125	0.003849	0.003879	0.002245	0.000641
24	780,842	3,140	0.004021	0.003943	0.002349	0.000677
25	779,670	3,154	0.004045	0.004133	0.002377	0.000708
26	669,409	3,169	0.004734	0.004559	0.002735	0.000736
27	644,261	3,184	0.004942	0.005020	0.002853	0.000763
28	668,283	3,217	0.004814	0.005071	0.002812	0.000810
29	592,344	3,249	0.005485	0.004748	0.003171	0.000858
30	779,884	3,282	0.004208	0.004781	0.002559	0.000909
31	419,800	3,314	0.007894	0.005580	0.004429	0.000963
32	608,199	3,347	0.005503	0.006456	0.003282	0.001021
33	507,593	3,513	0.006921	0.006614	0.004018	0.001115
34	493,407	3,678	0.007454	0.006529	0.004335	0.001216
35	573,950	3,844	0.006897	0.007281	0.004010	0.001323
36	484,691	4,009	0.008271	0.008668	0.004856	0.001440
37	399,857	4,175	0.010441	0.009303	0.006004	0.001586
38	499,488	4,237	0.008483	0.008857	0.005087	0.001692
39	410,585	4,299	0.010471	0.008761	0.006150	0.001830
40	557,256	4,362	0.007828	0.009894	0.004903	0.001978
41	235,990	4,424	0.018747	0.010859	0.010441	0.002135
42	394,795	4,886	0.012376	0.009968	0.007339	0.002302
43	717,705	4,894	0.006540	0.008270	0.004181	0.002445
44	287,011	4,902	0.017079	0.008965	0.008725	0.002595

TABLE 5

45	415,071	5,111	0.012314	0.013500	0.006800	0.002752
46	270,502	5,319	0.019663	0.019265	0.010003	0.002914
47	245,784	5,527	0.022486	0.021868	0.011293	0.003080
48	329,741	5,762	0.017474	0.019819	0.009374	0.003430
49	258,179	5,968	0.023232	0.016777	0.012018	0.003791
50	399,290	6,233	0.015610	0.018348	0.009011	0.004168
51	149,779	6,489	0.043190	0.025580	0.020913	0.004570
52	244,471	6,704	0.027422	0.032508	0.014490	0.005003
53	214,200	7,102	0.033156	0.032810	0.017122	0.005360
54	224,178	7,500	0.033456	0.029090	0.017475	0.005752
55	267,284	7,897	0.029545	0.031262	0.016064	0.006174
56	199,858	8,295	0.041504	0.042805	0.021382	0.006620
57	158,747	8,693	0.054780	0.053012	0.027261	0.007088
58	199,318	9,084	0.045475	0.049588	0.023824	0.007941
59	150,413	9,436	0.082734	0.037445	0.031637	0.008826
60	339,395	9,807	0.026866	0.036022	0.017855	0.009756
61	84,984	10,179	0.119775	0.054491	0.056882	0.010746
62	144,854	10,550	0.072933	0.077401	0.037675	0.011812
63	140,806	10,920	0.077664	0.082680	0.039973	0.012324
64	131,781	11,290	0.085685	0.070339	0.043700	0.012901
65	214,043	11,661	0.054480	0.064728	0.030850	0.013516
66	103,761	12,031	0.115949	0.085071	0.057224	0.014145
67	96,633	12,401	0.125347	0.118481	0.061557	0.014762
68	119,831	12,454	0.103930	0.126680	0.053418	0.016364
69	79,442	12,507	0.157436	0.100354	0.076984	0.017966
70	196,500	12,560	0.063919	0.074426	0.038363	0.019617
71	39,844	12,813	0.318560	0.112157	0.146268	0.021349
72	60,746	12,866	0.156862	0.199104	0.079763	0.023205
73	58,687	13,194	0.224820	0.230568	0.110727	0.027033
74	57,219	13,722	0.239615	0.166826	0.119418	0.031099
75	109,133	14,250	0.130575	0.150247	0.075716	0.035473
76	57,223	14,778	0.258253	0.250091	0.132486	0.040228
77	42,214	15,306	0.362581	0.306242	0.179638	0.045437
78	62,149	15,474	0.248982	0.284517	0.134055	0.049749
79	42,282	15,641	0.369921	0.327328	0.187989	0.054530
80	111,104	15,809	0.142290	0.148073	0.094863	0.058725
81	212,200	15,976	0.075287	0.074928	0.069513	0.065277
82	33,544	16,144	0.481278	0.483127	0.244886	0.071130
83	27,623	18,845	0.682221	0.681920	0.333706	0.076048
84	28,521	21,546	0.755443	0.755467	0.368840	0.085243
85	48,274	24,248	0.502299	0.502299	0.268050	0.092747

5.3 Comparación de los resultados.

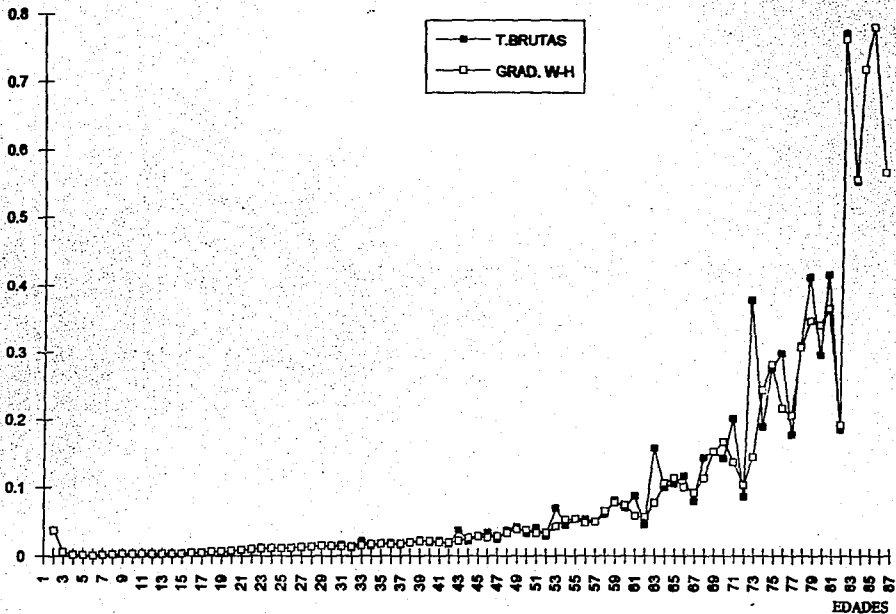
La comparación realizada en esta sección se llevará a cabo por medio de las gráficas de cada tabla graduada.

I Tasas brutas vs. tasas de la graduación Whittaker-Henderson.

Observando las gráficas comparativas tanto de hombres como de mujeres podemos notar que hasta edad 70 se obtuvo un grado de suavización relativamente bueno, los grandes brinco que se siguen detectando en las edades siguientes son provocadas en parte por el gran sesgo con las que fueron calculadas las tasas brutas debido a la información de la que provienen.

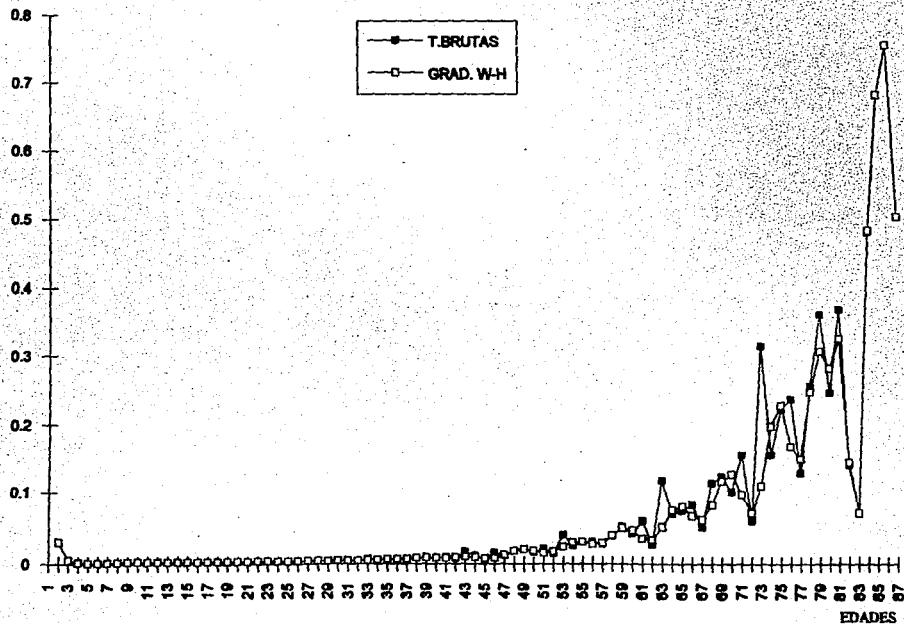
PROB. DE MUERTE

GRAFICA COMPARATIVA HOMBRES



PROB. DE MUERTE

GRAFICA COMPARATIVA MUJERES



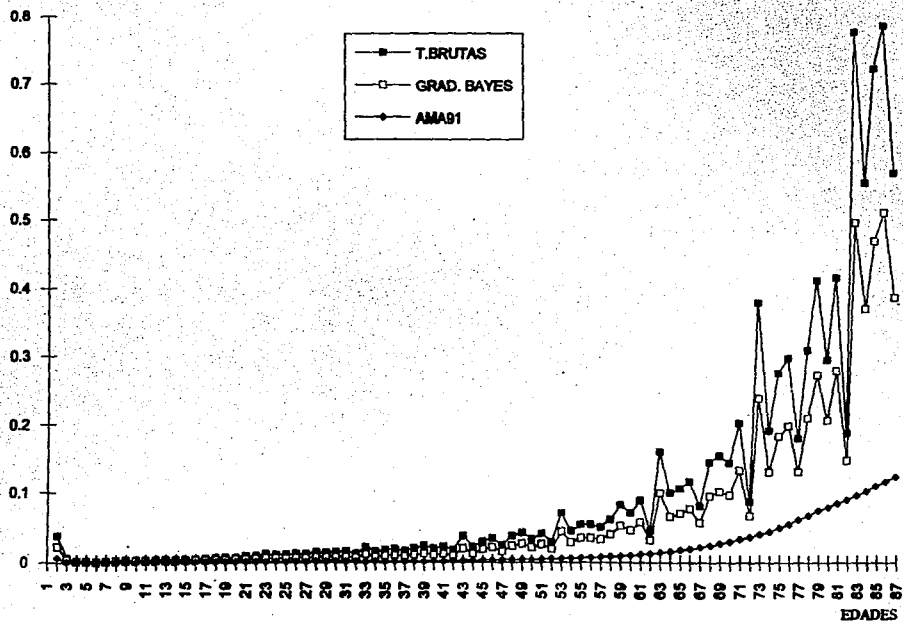
II Tasas brutas, tasas de la graduación de Bayes y tasas de la tabla de mortalidad AMA91.

En estos casos la tabla de mortalidad AMA91 nos sirvió como antecedente para aplicar la graduación de Bayes.

Nótese como nuestra curva graduada cae entre la de las tasas brutas y la de la AMA91.

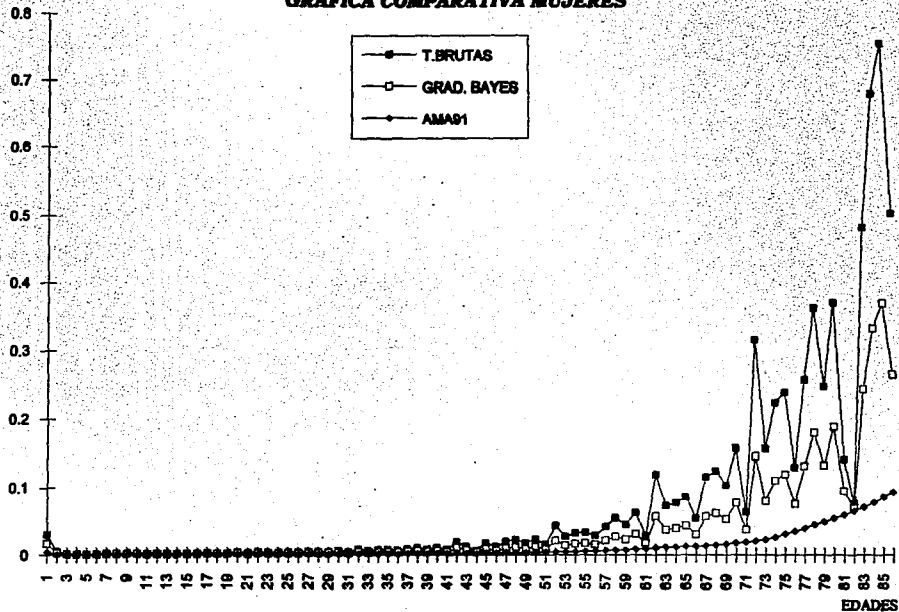
PROB. DE MUERTE

GRAFICA COMPARATIVA HOMBRES



PROB. DE MUERTE

GRAFICA COMPARATIVA MUJERES



III Tasas de la graduación de Whittaker-Henderson vs. tasas de la graduación de Bayes.

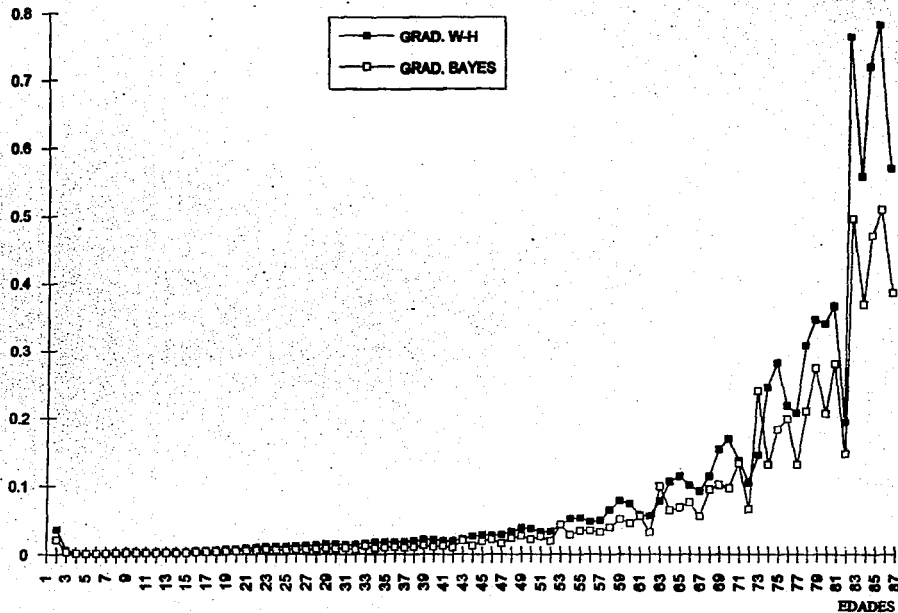
Podemos notar como en ambas gráficas (hombres y mujeres) la curva con un mayor grado de suavización es la obtenida por la técnica Bayesiana.

Esto podría deberse a la corrección que no provocó la inclusión de las tasas de la tabla de mortalidad AMA91.

Por lo que parecería más confiable, para este caso, la graduación obtenida por la técnica Bayesiana.

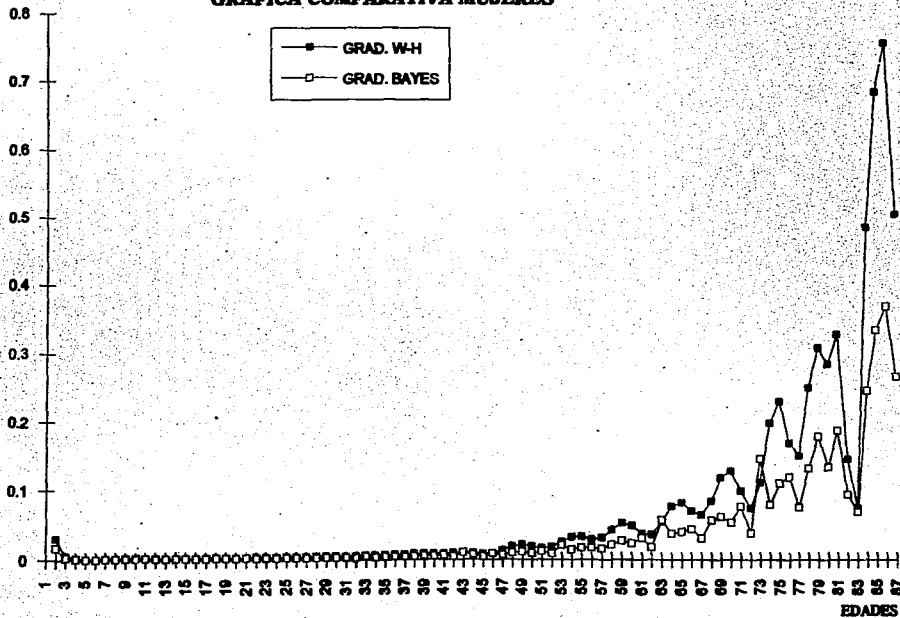
PROB. DE MUERTE

GRAFICA COMPARATIVA HOMBRES



PROB. DE MUERTE

GRAFICA COMPARATIVA MUJERES



CONCLUSIONES

Entre las conclusiones más relevantes a las que se llegó se tiene el hecho de que el trabajar con este tipo de teorías es prácticamente imposible sin la ayuda de una computadora que nos realice los cálculos, puesto que se manipulan matrices de dimensiones demasiado grandes para resolverlas manualmente.

Se remarcó la importancia de la graduación de tablas para fines prácticos y de tipo cálculo actuarial.

En cuanto a la comparación de las dos técnicas presentadas la más utilizada parece ser la de Whittaker-Henderson debido a que, como se explicó, es mucho más práctica y menos subjetiva que la Bayesiana. Por la información y bibliografía recolectada se observó que la graduación de Bayes es una de las técnicas más jóvenes que existen dentro de este campo en la actualidad, muchos autores auguran que esta será la que reemplace a la de Whittaker-Henderson, porque aunque más subjetiva, nos permite reflejar mejor nuestras creencias *a priori* o de interrelación entre las estimaciones iniciales.

De hecho resulta más sencillo desarrollar la teoría de Whittaker-Henderson y su ejemplo práctico que la teoría y ejemplo de la graduación de Bayes. Sin embargo la que suscribe la presente tesis es de la opinión de que si se cuenta con los elementos necesarios para desarrollar la graduación de Bayes se realice la graduación por medio de esta técnica que luce más acertada, ya que nos permite incluir la experiencia que se ha ido acumulando y no arrancar de cero como en el caso de Whittaker-Henderson, ejemplo claro de esto es la corrección que produjo incluir la experiencia de una tabla existente (AMA91) en la nueva que se construyó, ya que nos condujo a obtener una curva mucho más suave que con la otra técnica. Esto resulta de suma importancia dado que la información que se graduó contiene un alto grado de sesgo.

Por último es de interés notar que con los parámetros propuestos se obtuvo un mayor grado de ajuste que de suavización en las curvas graduadas del ejemplo práctico, (tanto por Whittaker-Henderson como por la Bayesiana), esto puede modificarse:

Primero, corrigiendo la información con la que se calcularon las tasas brutas (disminuir el sesgo en la información).

Segundo, experimentando con parámetros diferentes a los aquí propuestos.

BIBLIOGRAFIA.

1. BURDEN, R.L. Y J.D. FAIRES. *Análisis Numérico*. México: Grupo Editorial Iberoamérica, 1992.
2. CHAN, F.Y., L.K. CHAN Y M.H. YU. "A Generalization of Whittaker-Henderson Graduation", TSA XXXVI (1984), 183-211.
3. FREUND, J.E. Y R.E. WALPOLE. *Estadística Matemática con aplicaciones*, (4ta. edición). México: Prentice Hall, 1990.
4. FULKS, W. *Cálculo avanzado*. México: Limusa, 1986.
5. HICKMAN, J.C. Y R.B. MILLER. "Notes on Bayesian Graduation", TSA XXIX (1977), 1-21.
6. INEGI. *Censo de población general de los Estados Unidos Mexicanos*, 1990.
7. INEGI. *Defunciones en los Estados Unidos Mexicanos*, 1976-1992.
8. ITAM. *Tabla de mortalidad AMA91*.
9. JORDAN, C.W. *Life contingencies*, (2nd. edition). Chicago, Illinois: The Society of Actuaries, 1975.
10. KREYSZIG, E. *Introducción a la Estadística Matemática*. Principios y métodos. México: Limusa, 1989.
11. LONDON, D. *Graduation. The revision of estimates*. Winsted, Connecticut: ACTEX publications, 1985.
12. LOWRIE, W.B. "An Extension of the Whittaker-Henderson Method of Graduation", TSA XXXIV (1982), 329-72.
13. POLIVKA, R.P., S. PAKIN. *APL The language and its usage*. Englewood Cliffs, New Jersey: Prentice-Hall, Series in Automatic Computation, 1975.
14. PRESS, S.J. *Bayesian Statistics: Principles, Models and Applications*. U.S.A.: John Wiley & Sons, 1989.

15. SOUTHWORTH, R.W., S.L. DELEEUW. *Digital computation and numerical methods*. U.S.A.: Mc. Graw Hill, 1965.
16. SPIEGEL, R.M. *Estadística*, (2da. edición). México: Mc. Graw Hill, 1991.
17. WALPOLE, R.E., R.H. MYERS. *Probabilidad y Estadística*, (4a. edición). México: Mc. Graw Hill, 1992.

***Herramientas computacionales.
Paquetería utilizada.***

- Procesador de textos WinWord.
- Hoja de cálculo Excel.
- Lenguaje de programación APL.

ANEXOS

Anexo 1

Los métodos que se presentarán a continuación son útiles en la estimación de parámetros desconocidos de una población.

*Método de momentos*¹⁶.

El método de momentos consiste en igualar los primeros momentos de una población a los momentos correspondientes de una muestra, con lo cual se obtienen tantas ecuaciones, según se necesiten, para resolver y obtener los parámetros desconocidos de la población.

Definición.- El k -ésimo momento de las muestra de un conjunto de observaciones x_1, x_2, \dots, x_n es la medida de sus k -ésimas potencias y se presenta por medio de m'_k ; en forma simbólica, se tiene:

$$m'_k = \frac{\sum_{i=1}^n x_i^k}{n}$$

Por lo tanto, el método de momentos consiste en solucionar el sistema de ecuaciones:

$$m'_k = \mu'_k \quad k = 1, 2, \dots, p$$

para obtener los p parámetros de la población.

*Método de máxima verosimilitud*¹⁷.

Este método tiene la ventaja de que los estimadores que produce son suficientes, siempre que estos existan, y que son asintóticamente insesgados de varianza mínima.

¹⁶Ver [3] de la bibliografía.

¹⁷Ver [3] de la bibliografía.

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

La característica esencial del método de máxima verosimilitud es que primero observamos los valores de una muestra aleatoria y después elegimos como nuestra estimación del parámetro desconocido de la población, el valor para el cual la probabilidad de obtener los datos observados es un máximo.

Definición.- Si $x_1, x_2, \dots, y x_n$ son los valores de una muestra tomada al azar de una población con el parámetro Θ , la función de verosimilitud de la muestra está dada por:

$$L(\Theta) = f(x_1, x_2, \dots, x_n; \Theta)$$

para valores de Θ contenidos en un dominio dado. Aquí $f(x_1, x_2, \dots, x_n; \Theta)$ es el valor de la distribución de probabilidad conjunta o función de densidad conjunta de las variables aleatorias $X_1, X_2, \dots, y X_n$ es el punto de la muestra observado. Por consiguiente, el método de máxima verosimilitud consiste en maximizar la función de verosimilitud con respecto a Θ y nos referimos al valor de Θ que maximiza la función de probabilidad como la estimación de máxima verosimilitud de Θ .

Anexo 2

Aproximación de la distribución binomial mediante la distribución normal¹⁸.

La distribución normal se puede obtener de una aproximación de la distribución binomial cuando n es grande. Recordemos que la distribución binomial tiene la siguiente función de probabilidades:

$$f(x) = \binom{n}{x} p^x q^{n-x} \quad (x = 0, 1, \dots, n), \quad \dots(1)$$

con media:

$$\mu = np \quad \dots(2)$$

y varianza:

$$\sigma^2 = npq \quad \dots(3)$$

Teorema. - Sea $0 < p < 1$. Entonces, para n grande, la distribución binomial se puede aproximar por medio de la distribución normal con media dada por (2) y varianza dada por (3), esto es:

$$f(x) \cong f^*(x) \quad (x = 0, 1, \dots, n),$$

donde:

$$f^*(x) = \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-z^2/2}, \quad z = \frac{x - np}{\sqrt{npq}}$$

es la densidad de la distribución normal, y el símbolo \cong (léase asintóticamente igual) significa que la razón de ambos miembros se aproxima a 1 cuando n tiende a ∞ .

¹⁸Ver [10] de la bibliografía.

Anexo 3.

Método de factorización de Choleski¹⁹.

Recordemos que cuando se quiere resolver un sistema lineal de la forma $Ax=b$ podemos hacerlo por eliminación Gaussiana o bien factorizando la matriz A en términos de una matriz inferior L y de una matriz triangular superior U (llamado método LU).

El método de Choleski es precisamente una técnica para llevar a cabo esta factorización.

El procedimiento de factorización se describe en el siguiente algoritmo:

Algoritmo de Choleski.

Para factorizar una matriz $n \times n$ positiva definida A como LL^t donde L es triangular inferior:

ENTRADA la dimensión n ; los elementos a_{ij} , $1 \leq i, j \leq n$ de A .

SALIDA los elementos l_{ij} , $1 \leq j \leq i, 1 \leq i \leq n$ de L . (Los elementos de $U = L^t$ son $u_{ij} = l_{ji}$, $i \leq j \leq n, 1 \leq i \leq n$).

PASO 1 Tomar $l_{11} = \sqrt{a_{11}}$.

PASO 2 Para $j=2, \dots, n$ tomar $l_{j1} = \frac{a_{j1}}{l_{11}}$.

PASO 3 Para $i=2, \dots, n-1$ seguir los Pasos 4 y 5.

PASO 4 Tomar $l_{ii} = \left[a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right]^{1/2}$.

¹⁹Ver [1] de la bibliografía.

PASO 5 Para $j=i+1, \dots, n$

$$\text{tomar } l_{ji} = \frac{1}{l_{ii}} \left[a_{ji} - \sum_{k=i}^{j-1} l_{jk} l_{ik} \right].$$

PASO 6 Tomar $l_{nn} = \left[a_{nn} - \sum_{k=1}^{n-1} l_{nk}^2 \right]^{1/2}$.

PASO 7 SALIDA (l_{ij} para $j=1, \dots, i$ e $i=1, \dots, n$);

PARAR.

PROGRAMAS EN APL.

**PROGRAMA QUE REALIZA LA GRADUACION DE
WHITTAKER-HENDERSON.**

```
[0] W+n WEIGHTS u;Q
[1] A+((1+pn)x(ux1-u)=0)+((ux1-u)≠0)xn+(ux1-u)+(ux1-u)=0
[2] Q+(((pA),1)pA),((pA),(pA))p0
[3] W+((pA),(pA))pQ
```

```
[0] Q+Z KZ N
[1] +(0=2|Z)pL1
[2] Q+((N-Z),N)p((-1*1(Z+1))x(0,1Z)|Z),(N-Z)p0
[3] +0
[4] L1:Q+((N-Z),N)p((-1*1(Z+1))x(0,1Z)|Z),(N-Z)p0
```

```
[0] V+n WHITT E;Q
[1] u+E+n
[2] W+n WEIGHTS u
[3] 'TECLEE EL ORDEN Z DE LAS DIFERENCIAS FINITAS'
[4] K+□ KZpn
[5] nH+l(+/A)+pn
[6] H+0.5
[7] X+(1-H)x(W+.xu)□((1-H)xW+Hx(□K)+.xK)
[8] nX+(W+.xu)□(W+Hx(□K)+.xK)
[9] V+ 0 6 √(((pn),1)pX)
```

FALLA DE ORIGEN
FALLA DE ORIGEN

PROGRAMA QUE REALIZA LA GRADUACION BAYESIANA DE KIMELDORF-JONES.

```

[ 0 ] V←Bayes
[ 1 ] 'Introduzca el vector de muertes...'
[ 2 ] E←□
[ 3 ] 'Introduzca el vector de poblacion total...'
[ 4 ] n←□
[ 5 ] 'Introduzca el vector de estimaciones ya existentes...'
[ 6 ] m←□
[ 7 ] 'Introduzca el valor inicial del coeficiente de correlacion...'
[ 8 ] r←□
[ 9 ] 'Limitado a la vecindad entre cuantos elementos??...'
[10 ] ccv←□
[11 ] u←E+n
[12 ] I←(1pn)°.-1pn
[13 ] aux←(m1-m)+n
[14 ] p←(+/aux)+p aux
[15 ] a←(p*2)xr*0,1ccv
[16 ] a1←((pn),(pn))pa,((pn)-pa)p0,0
[17 ] a2←a1x(1pn)°.-1pn
[18 ] a3←(a1)x(1pn)°.-21pn
[19 ] A←a2+a3
[20 ] ac1←a2[143;143]+a3[143;143]
[21 ] ac2←a2[143;(43+143)]+a3[143;(43+143)]
[22 ] ac3←a2[(43+143);143]+a3[(43+143);143]
[23 ] ac4←a2[(43+143);(43+143)]+a3[(43+143);(43+143)]
[24 ] i←1
[25 ] loop:A[i;1]+p*2
[26 ] +((pn)2i+1+1)/loop
[27 ] B←(((pn),(pn))paux)x(1pn)°.-1pn
[28 ] V←u+(A+xB)Bm-u

```