

00321



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

77

FACULTAD DE CIENCIAS

"ANALISIS Y PROPUESTA DE METODOS DE AJUSTE PARA TABLAS DE MORTALIDAD"

T E S I S  
QUE PARA OBTENER EL TITULO DE:  
A C T U A R I O  
P R E S E N T A:  
CESAR PONZANELLI VELAZQUEZ



FACULTAD DE CIENCIAS UNAM

DIRECTOR DE TESIS: ACT. EDGAR ARTURO TARANGO MENDEZ

2003

TESIS CON FALLA DE ORIGEN



FACULTAD DE CIENCIAS SECCION ESCOLAR



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 esta 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.

# **PAGINACION DISCONTINUA**



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**DRA. MARÍA DE LOURDES ESTEVA PERALTA**  
**Jefa de la División de Estudios Profesionales de la**  
**Facultad de Ciencias**  
**Presente**

Comunicamos a usted que hemos revisado el trabajo escrito: "Análisis y propuesta de Métodos de Ajuste para Tablas de Mortalidad" realizado por César Ponzanelli Velázquez

con número de cuenta 9251604-7, quien cubrió los créditos de la carrera de: Actuaría.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis

Propietario Act. Edgar Arturo Terango Méndez

Propietario Act. Fernando Eleazar Vanegas Chávez

Propietario Act. Ricardo Humberto Sevilla Aguilar

Suplente Act. Alfredo Arceo Franco

Suplente Act. Miguel Angel Torres Ramírez

*[Handwritten signatures and initials]*

Consejo Departamental de Matemáticas.

*[Handwritten signature]*  
M. en C. José Antonio *[Handwritten]*  
FACULTAD DE CIENCIAS  
CONSEJO DEPARTAMENTAL  
DE  
MATEMÁTICAS

Un especial agradecimiento a Edgar Tarango por que gracias a su apoyo y confianza fue posible la realización de esta meta.

Dedicada a mis padres Tomás y Julieta, a mi abuelita María de la Paz, a mis hermanas Julieta, Rossana y Gloconda, a mi cuñado Luis Alonso y a mi novia Raquel. A todos ellos por su ayuda, interés y compañía.

A mi sobrino por su afecto, sinceridad y noble espíritu que fueron un estímulo para seguir adelante.

A la memoria de María Javier Hernández y Enrique Monedero, por tomarse el tiempo de enseñarme importantes lecciones que ningún temario debía abarcar.

Mi agradecimiento a todas las personas que contribuyeron a la realización de este objetivo. A la Jefa en mi oficina por darme todos los permisos necesarios, así como a la UNAM por brindarme esta oportunidad.

A todos ustedes, mil y un gracias.

# Índice.

Introducción. ....	1
I. Primeros Conceptos. ....	1
1.1 Antecedentes Históricos en México. ....	2
1.2 De los Datos Muestrales a las Tablas de Mortalidad. ....	5
1.2.1 Búsqueda de Errores en los Datos. ....	5
1.2.2 Elaboración de las Tasas Brutas de Mortalidad. ....	6
1.2.3 Suavizar las Tasas de Mortalidad. ....	6
II. Elementos del Cálculo Actuarial para Vida. ....	7
2.1 Tablas de mortalidad. ....	7
2.1.1 Elementos de la Tabla de Mortalidad. ....	7
2.1.2 Tabla de Mortalidad Experiencia Mexicana. ....	9
2.1.3 Cálculo de los Elementos de la Tabla de Mortalidad. ....	13
2.1.4 Equivalencias entre los Elementos de la Tabla de Mortalidad. ....	18
2.2 Función de Mortalidad. ....	21
2.2.1 Fuerza de Mortalidad $\mu_x$ . ....	22
2.2.2 Función de Supervivencia $s(x)$ . ....	23
2.2.3 Función de Mortalidad $F(x)$ . ....	27
2.2.4 Cálculo de los Elementos de la Tabla de Mortalidad con $\mu_x$ . ....	29
2.3 Variables Aleatorias para Mortalidad. ....	32
2.3.1 Variable Aleatoria $X \sim$ Edad de Muerte. ....	32
2.3.2 Variable Aleatoria $T_{(x)} \sim$ Tiempo Restante de Vida. ....	34
2.3.3 Variable Aleatoria $K_{(x)} \sim$ Años Enteros Restantes de Vida. ....	36
2.4 Otras Funciones Vitales. ....	38
2.4.1 Esperanza Completa de Vida. ....	38
2.4.2 Tiempo Medio de Vida $m(x)$ . ....	39
2.4.3 Años Persona Vividos $L_x$ . ....	40
2.4.4 Total Años Persona Vividos $T_x$ . ....	40
2.4.5 Tasa Central de Mortalidad $m_x$ . ....	41
2.5 Tablas Selectas de Mortalidad. ....	41
2.5.1 Probabilidades Selectas. ....	43
2.5.2 Ejemplos de Cálculo de Probabilidades Selectas. ....	45
2.5.3 Propuesta de Función Selecta de supervivencia. ....	47

III. Método de Ajuste AMIS-AMA. ....	49
3.1 Primera Ley de Makeham. ....	49
3.2 Cálculo de $p_x$ Utilizando la Ley de Makeham. ....	50
3.3 Determinando los Parámetros $s$ , $g$ , y $C$ . ....	50
3.4 Aplicación. ....	53
3.5 Resultados. ....	55
3.6 Alcances y Limitaciones. ....	57
IV. Método de ajuste Propuesta. ....	58
4.1 Polinomios. ....	58
4.2 Método de Ajuste. ....	63
4.3 Aplicación. ....	77
4.4 Resultados. ....	81
4.5 Alcances y Limitaciones. ....	84
Conclusiones. ....	85
Anexo. Probabilidad. ....	87
I. Enfoque Clásico o Probabilidad a Priori. ....	89
II. Enfoque Frecuentista o Probabilidad a Posteriori. ....	92
III. Probabilidad Axiomática. ....	95
3.1 Teoría de Conjuntos. ....	96
3.2 Definición de Espacio muestral y evento. ....	102
3.3 Definición de Probabilidad. ....	108
3.4 Espacios Muestrales Finitos. ....	116
3.5 Probabilidad Condicional e Independencia. ....	124
Bibliografía. ....	137

## Introducción.

El presente trabajo tiene dos objetivos principales: el primero es el de examinar métodos de ajuste para tablas de mortalidad, así como proponer un método alternativo; el segundo es explicar los principales elementos y funciones que intervienen en el desarrollo de una tabla de mortalidad y en el cálculo actuarial para vida individual.

La *Tabla de Mortalidad* ha sido el instrumento tradicional para almacenar las expectativas de mortalidad para una población. No obstante que ésta es una fuente práctica de información, no es lo suficientemente completa para permitir el cálculo de cualquier prima de Seguro de Vida. Existen diversos planes de seguros que requieren de medidas instantáneas de mortalidad lo cual solo es posible con funciones continuas. Además las fórmulas idóneas para el cálculo de primas de seguros plantean el manejo de funciones continuas y no un conjunto de datos discretos como lo es la tabla de mortalidad.

Otra importante razón para el desarrollo de modelos matemáticos relacionados con la mortalidad, es que durante la elaboración de una tabla de mortalidad es común encontrarse con desviaciones no reales sobre las probabilidades de muerte. Tales desviaciones se identifican por estar asociadas a fenómenos socio-económicos o a muestras no significativas en ciertas edades; sin embargo pueden ser corregidas ajustando los datos mediante alguna curva para después tomar los valores de la curva como verdaderos.

Es decir, por las ventajas que plantea una función de mortalidad, es necesario el desarrollo y estudio de modelos que puedan servir igual o mejor que los esquemas que derivan en una tradicional tabla de mortalidad.

Las primeras nociones sobre tablas de mortalidad y los antecedentes históricos en México se encuentran en el capítulo 1.

El capítulo 2 es el más extenso, pues contiene las definiciones, funciones y equivalencias elementales necesarias para el desarrollo de la tabla de mortalidad y el cálculo actuarial para vida individual. Si estos temas no son del dominio del lector, se aconseja repasar el contenido de este capítulo antes de revisar los capítulos restantes.

En los capítulos 3 y 4 se examinarán métodos de ajuste para la representación de los datos de la tabla de mortalidad, basados en modelos matemáticos.

En el capítulo 3 se desarrolla el método de ajuste utilizado en el documento "Tablas de Mortalidad México 2000", realizado por la Asociación Mexicana de Instituciones de Seguros, A.C. (AMIS) y la Asociación Mexicana de Actuarios, A.C. (AMA). En dicho documento se elaboran tablas de mortalidad con la Información aportada por las instituciones de seguros, además de presentar un método de ajuste para suavizar tales datos a partir de la ley de Makeham.

El capítulo 4 contiene la propuesta de un método alterno para ajustar tablas de mortalidad con polinomios, que es el tema central del presente trabajo.

Es necesario aclarar que a las tablas de mortalidad y a las probabilidades de muerte, se les dará un manejo desde el punto de vista del cálculo actuarial para vida, no obstante son tema común de demografía.

## I. Primeros Conceptos.

En general, el Seguro es un contrato en el que se compromete el pago de una cantidad mayor de dinero denominada Suma Asegurada, para solventar la pérdida económica provocada por la posible ocurrencia de un evento económicamente desfavorable. El costo por tal protección se denomina Prima.

En particular el Seguro de vida es el que cubre la pérdida económica a los Beneficiarios, provocada por la muerte o supervivencia del Asegurado dentro de un periodo de tiempo.

A manera de ejemplo, se pudiera pensar en un plan de seguro el cual paga una suma asegurada repartida entre esposa e hijos en caso que fallezca el padre de familia. Otro ejemplo es un plan que pague una suma asegurada necesaria para garantizar la educación superior de un hijo, en caso que este sobreviva a la edad suficiente para entrar a la universidad.

Entonces se hace evidente que para poder calcular las primas para los Seguros de Vida, es necesario contar con fuentes de información que resuman de alguna manera la expectativa que tiene una persona de morir o sobrevivir dentro de un periodo de tiempo. El instrumento tradicional para almacenar y acceder a dicha información ha sido la *Tabla de Mortalidad*.

Se puede definir a la tabla de mortalidad, también conocida como tabla de vida, a la estadística que resume la experiencia de una población ante el evento muerte.

Es la forma en que se distribuyen las expectativas de supervivencia (o muerte) en cada edad, lo que determina los valores y da forma a la tabla de mortalidad. Por tal razón, cada población tiene una tabla de mortalidad diferente a las demás poblaciones y más aún, subconjuntos particulares de una población pueden llegar a presentar tasas de mortalidad significativamente diferentes a la del resto de la población. Así, podemos tener tablas de mortalidad por hombre y mujeres, por fumadores y no fumadores, etc.

Las tablas de mortalidad no son estáticas. Conforme una población cambia con el paso del tiempo, cambian sus expectativas de supervivencia y por lo tanto su tabla de mortalidad.

Se dividen las tablas de mortalidad en tablas a edad alcanzada o tablas selecta y última. Una tabla de mortalidad selecta y última, hace referencia a un subconjunto de la población con expectativas favorables de vida, la cual fue seleccionada mediante exámenes médicos. En una tabla de mortalidad a edad alcanzada la población que representa no fue seleccionada mediante algún criterio.

También se clasifican las tablas de mortalidad en completas y abreviadas. Una tabla de mortalidad completa contiene las tasas de mortalidad para todos los posibles grupos de edades de una población. Una tabla de mortalidad abreviada contiene las tasas de mortalidad para solo algunos grupos de edades de una población.

Por último y dándole un mejor énfasis, es conviene describir brevemente que son las tablas de decremento múltiple. Mientras una tabla de mortalidad se construye bajo la suposición que sus integrantes son un grupo cerrado del cual solo existen salidas por muerte, en una tabla de decremento múltiple existe la posibilidad de abandonar el grupo por dos o más causas, por ejemplo mortalidad, invalidez, retiro, etc. De esta manera, las tablas de decremento múltiple son un conjunto de dos o más tablas, donde cada tabla expresa las probabilidades de abandonar el grupo por cada causa en particular.

En la práctica, una tabla de mortalidad se elabora a partir de tablas de decremento múltiple, pues como es de suponerse, en la vida real no existen grupos cerrados en los cuales sus integrantes solo lo abandonen por muerte. Se observa un grupo control, como lo es por ejemplo la colectividad asegurada de las compañías aseguradoras y se registran las salidas del grupo así como fecha y causa. Aquellas que sean por muerte serán la base para estimar la tabla de mortalidad. Además en el seguro de vida existen beneficios adicionales para los cuales la Prima se puede calcular con tablas de decremento que además de muerte manejen invalidez y morbilidad.

## 1.1 Antecedentes Históricos en México.

En un mercado competido, el actuario debe estimar la prima mínima necesaria a cobrar sin tener pérdidas para poder garantizar a la Aseguradora un volumen adecuado de ventas y no quedar fuera de mercado.

Esta meta requiere de cálculos precisos y arriesgados, los cuales solo son posibles con tablas de mortalidad que representen fielmente a la población que se pretende asegurar. Por esto, en una sociedad socio-económicamente dinámica una tabla de mortalidad útil en una fecha resulta inadecuada en pocos años.

Por lo tanto, la competencia entre Aseguradoras origina la necesidad de desarrollar constantemente nuevas tablas de mortalidad, con lo cual se obtiene un historial de tablas a través del tiempo.

La tabla de mortalidad conocida como "Experiencia Americana" fue la primera en ser utilizada por Aseguradoras en nuestro país. Esta tabla fue publicada entre 1864 y 1867 basándose en 20 años de experiencia de la "Mutual Life Insurance Co. of New York".

Conforme maduraba el sector asegurador en nuestro país, se hizo evidente que la tabla "Experiencia Americana" no representaba adecuadamente a la mortalidad mexicana. El principal problema de esta tabla era que tenía tasas de mortalidad muy conservadoras sobre todo a edades jóvenes. Nació así la necesidad de desarrollar una tabla más reciente, creada a partir de las experiencias observadas sobre la población que se pretendía asegurar.

EL primer intento por actualizar la tabla de mortalidad por una más precisa se hizo para el Seguro de Vida en Grupo, cuando se logró obtener la autorización de las autoridades para el uso de la tabla C.S.O. 1941 (The Commissioners Standard Ordinary). No obstante se trataba de una tabla más reciente, se presentaban devoluciones de primas de un 25% en promedio por concepto de dividendo por mortalidad <sup>(1)</sup>.

La primera tabla de vida desarrollada en experiencias mexicanas fue la denominada "Experiencia Mexicana 62-67". El desarrollo de esta tabla convino con grandes contrastes en el sector asegurador de nuestro país, pues al fin se contaba con una estadística representativa de la población mexicana que permitía cálculos más veraces y precisos.

---

<sup>(1)</sup> El dividendo por mortalidad es una práctica frecuente en las pólizas de vida grupo o colectivo. Consiste en hacer partícipe al contratante de las utilidades generadas por una siniestralidad favorable, esto es cuando la mortalidad estimada fue mayor a la experimentada. El dividendo se paga al aniversario de la póliza.

El artículo 67 de la Ley General de Instituciones de Seguros indicaba el uso de la tabla "Experiencia Americana" o una más conservadora para la constitución de la reserva matemática. Sin embargo, las compañías Aseguradoras en el continuo afán de ser competitivas, para obtener tarifas más moderadas comenzaron a utilizar la tabla "Experiencia Mexicana 62-67", puesto que con la tabla "Experiencia Americana" se obtenían tarifas muy elevadas. Por consecuencia, muchas compañías Aseguradoras comenzaron a tarificar y proyectar con una tabla, y reportar reservas matemáticas "legales" con otra.

Otras compañías conscientes que la "Experiencia Americana" era una tabla muy conservadora, castigaban el cálculo de sus tarifas en el momento de incluir los gastos. Esto ocasionaba cálculos fuera de toda base real, con el inminente riesgo de calcular Primas por debajo del costo.

Esta situación propició una gran diversidad en los precios de productos semejantes entre Aseguradoras.

Otro problema resultado del caduco uso de la tabla Experiencia Americana se presentaba con los rescates de pólizas, sobre todo en efectivo. Los valores garantizados de una póliza se deben calcular en función de las reservas matemáticas proyectadas a través del tiempo, estimaciones que podían ser muy elevadas si se utilizaba la tabla "Experiencia Americana".

Como se podía esperar, la Ley General de Instituciones y Sociedades Mutualistas de Seguros no tardó en cambiar para incorporar en sus estatutos el uso de la nueva tabla.

Posteriormente, solo para el seguro de grupo se emitió la tabla experiencia mexicana 73-83. Después salió la tabla Experiencia Mexicana 82-89 solo para seguro de vida individual. Más reciente aún es la tabla Experiencia mexicana 91-98 elaborada tanto para seguro de vida individual (CNSF 2000-I (1991-1998)) como vida grupo (CNSF 2000-G (1991-1998)). Esta última tabla es la legalmente en curso, a partir del 1° de abril del 2000 según lo dispuesto por el acuerdo publicado en el Diario Oficial el viernes 31 de diciembre de 1999 entre la Comisión Nacional de Seguros y Fianzas (CNSF) y la Secretaría de Hacienda y Crédito Público (SHCP).

## 1.2 De los Datos Muestrales a las Tablas de Mortalidad.

Podemos diferenciar tres importantes pasos en el proceso de elaborar tablas de mortalidad a partir de los datos observados:

### 1.2.1 Búsqueda de Errores en los Datos.

Dependiendo del origen de los datos, se pueden presentar errores al momento de inscribir o recopilar la información. Una causa es por ejemplo, errores de registro por parte de las compañías Aseguradoras en los formatos establecidos por la CNSF. Actualmente para evitar este problema la CNSF impone a las Aseguradoras el uso de programas para registro de datos, los cuales validan la información desde el momento de la captura. Sin embargo en el pasado no se contaba con tal herramienta por lo que las últimas tablas de mortalidad se debieron elaborar con datos que presentaban errores evidentes.

El método para detectar errores en los datos es principalmente elaborar diversas tablas de mortalidad preliminares para cada forma útil de segmentar los datos: por compañía Aseguradora, por año de registro, por sexo, por condición de fumador, por antigüedad, etc. Los datos que presenten gráficamente irregularidades atípicas deberán ser analizados y de ser posible desechados. La elaboración de estos procedimientos es posible mediante el uso de manejadores de bases de datos.

En el documento "Análisis Exploratorio de los Datos del Seguro de Vida 1991-1998" elaborado por la CNSF se describe en forma detallada el procedimiento para segmentar y corregir los datos. Además se describe la base de datos elaborada para tal propósito que puede ser implementada por cualquiera que le interese.

## 1.2.2 Elaboración de las Tasas Brutas de Mortalidad.

Una vez que la base de datos se encuentra depurada, se puede proceder con la elaboración de las tasas brutas de mortalidad. Es posible que este paso requiera del manejo de tablas de decremento múltiple que no son tema del presente trabajo.

En los documentos "Tablas de Mortalidad México 2000" del cual previamente se hizo referencia y "Modelos Estadísticos de Mortalidad Análisis de Datos 1991-1998", éste último realizado por la CNSF, se detallan procedimientos para determinar las tasas de mortalidad brutas.

## 1.2.3 Suavizar las Tasas de Mortalidad.

El último paso consiste en eliminar las desviaciones significativas para suavizar la tabla de mortalidad; este hecho no solo tiene relevancia para eliminar los datos abruptos de las tasas previamente obtenidas si no además para obtener una curva que cumpla con todos los requerimientos de una función clase  $C^1$  <sup>(2)</sup>.

Esto se logra ajustando las tasas obtenidas mediante una curva que describa la tendencia de los datos. Esto es especialmente necesario en edades elevadas en las que se cuenta tasas de mortalidad muy grandes y con pocos datos muestrales, lo que ocasiona una gran dispersión en los datos.

Este proceso recibe el nombre de *ajuste o graduación por fórmula matemática de tablas de mortalidad* y es tema principal del presente trabajo.

---

<sup>(2)</sup> Una función clase  $C^1$  es aquella que tiene primera derivada continua.

---

## **II. Elementos del Cálculo Actuarial para Vida.**

El desarrollo del cálculo actuarial para el seguro de vida es posible gracias al hecho de que la distribución de muerte en una población no ocurre al azar, si no que está estrechamente relacionada con la edad; es decir, para cada persona es posible asociar en función a su edad una probabilidad de muerte y por consecuencia, también una probabilidad de supervivencia.

Estas probabilidades son el pilar del cálculo actuarial para vida, cuyo principio es el de estimar el importe de los derechos u obligaciones generados como consecuencia de la muerte o supervivencia del asegurado. Por tal razón, si el lector no se encuentra familiarizado con la teoría de probabilidades, se le recomienda consulte el anexo en la parte final del texto donde se encuentra un pequeño resumen con los puntos más necesarios al respecto.

Aunque hemos hablado indistintamente de tasas de mortalidad y probabilidades de muerte, es necesario aclarar que no son lo mismo. Mientras una tasa es un promedio unitario casi siempre respecto al tiempo, una probabilidad es una función con características muy específicas que asocia a un evento con la frecuencia con la que puede ocurrir dicho evento.

### **2.1 Tablas de Mortalidad.**

#### **2.1.1 Elementos de la Tabla de Mortalidad.**

Sin contradecir la definición del capítulo 1, podemos decir que una tabla de mortalidad es el ordenamiento de datos que contiene implícita o explícitamente, las probabilidades de muerte para cada edad en la que se segmentó la población. Por esto, es de esperarse que la mayoría de los elementos de la tabla de mortalidad sean probabilidades o conceptos directamente derivados de estas.

La mayoría de los elementos actuariales cuentan con una nomenclatura ya establecida. Es importante conservarla para que podamos darnos a entender en un ámbito general.

La siguiente es una lista de los principales elementos de la tabla de mortalidad. El cálculo y las equivalencias entre estos se deja más adelante:

- (x). Donde observemos este símbolo, debemos sustituirlo mentalmente por "una persona de edad x".
- ${}_t p_x$  denota a la probabilidad que tiene (x) de llegar con vida a edad  $x + t$ . Cuando t es igual 1 se omite escribirla y la notación se reduce a  $p_x$ .
- ${}_t q_x$  denota a la probabilidad que tiene (x) de morir antes de tener la edad  $x + t$ . Cuando t es igual 1 se omite escribirla y la notación queda como  $q_x$ .
- ${}_n|_m q_x$  denota a la probabilidad que tiene (x) de llegar con vida a edad  $x + n$  y además morir antes de tener la edad  $x + n + m$ . Cuando m es igual 1 se omite escribirla y la notación se reduce a  ${}_n|q_x$ . Si el diferimiento es nulo, es decir  $n = 0$ , se tiene que la probabilidad diferida se convierte en la probabilidad ordinaria de muerte  ${}_m q_x$ .
- $l_0$  se conoce como radix y es el número hipotético de recién nacidos con los que empieza la tabla. Es un número arbitrario y subjetivo, el cual solo sirve para calcular el esperado número de sobrevivientes a edad x ( $l_x$ ).
- $l_x$  es el número esperado de sobrevivientes a la edad x, partiendo de un grupo con  $l_0$  recién nacidos. Es decir,  $l_x$  es el teórico número de sobrevivientes a edad x, construido a partir de un radix inicial y de la probabilidad que tiene un recién nacido para llegar edad x.  
En las tablas abreviadas que no contienen los primeros años de vida, por ejemplo las que comienzan a los 12 años, no existe el radix  $l_0$ . Sin embargo el cálculo de  $l_x$ , es posible asignando un número arbitrario  $l_{inicial}$  al primer grupo de edades de la tabla.
- ${}_n d_x$  es el número esperado de muertes ocurridas entre las edades x y  $x + n$ . Cuando n es igual 1 se omite escribirla y la notación queda como  $d_x$ .

La probabilidad  ${}_n|_m q_x$  no es una función elemental y nunca aparece explícitamente en tablas de mortalidad. Sin embargo se incluye en esta sección, pues tiene una estrecha relación y similitud con las demás funciones de la tabla de mortalidad.

## 2.1.2 Tabla de Mortalidad Experiencia Mexicana.

La probabilidad anual de muerte  $q_x$  y la fuerza de mortalidad  $\mu_x$  cual definiremos más adelante, son los mejores elementos para caracterizar y comparar la mortalidad de una población.

Sin perder las características particulares de cada población, todas las tablas de mortalidad guardan el mismo esquema general. Durante el primer año de vida se registran probabilidades de muerte muy altas, las cuales disminuyen conforme se incrementa la edad hasta llegar a un valor mínimo. Posteriormente a este punto, las probabilidades de muerte crecen proporcionalmente conforme la edad se incrementa. Tal situación continúa sin cambios hasta que se extingue el grupo, punto en el que se marca el fin de la tabla de mortalidad.

De este comportamiento observamos dos principales factores de mortalidad, uno que es inversamente proporcional a la edad y otro que es directamente proporcional a la edad; el primero pudiera deberse a un proceso de adaptación del individuo al medio, en el que accidentes, problemas congénitos, enfermedades infecciosas o deficiencias del sistema inmune son las primeras causas de muerte; el segundo factor pudiera ser causa de un incremento a la exposición a factores de riesgo y por el envejecimiento natural del organismo.

La edad en la que se extingue el grupo se denota como  $w$  y se define de la siguiente manera:

$$w = \{\text{mínimo } \{x\} \mid x_{p_0} = 0\} = \{\text{mínimo } \{x\} \mid l_x = 0\}.$$

El esquema de mortalidad antes mencionado lo podemos verificar en la tabla "CNSF 2000 - I (1991-1998)" y en la gráfica que le sigue. Puesto que la "CNSF 2000 - I (1991-1998)" es una tabla abreviada y no contiene los valores entre cero y 11 años, estos se estimaron de manera proporcional utilizando de base otras tablas completas:

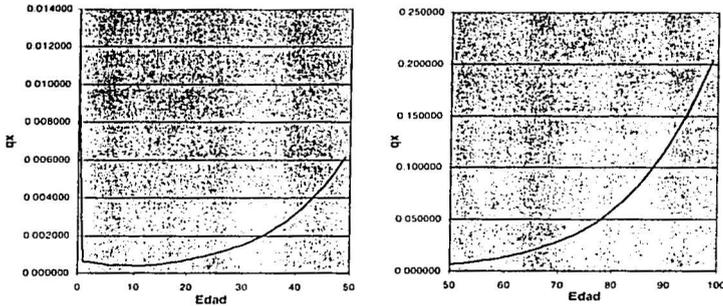
Tabla de Mortalidad experiencia mexicana 91 - 98 *CNSF 2000 - I (1991-1998)*.					
Edad	$q_x$	$p_x$	${}_1p_0$	$l_x$	$d_x$
0	0.012816	0.987184	1.000000	100,000	1,282
1	0.000679	0.999321	0.987184	98,718	67
2	0.000611	0.999389	0.986514	98,651	60
3	0.000554	0.999446	0.985911	98,591	55
4	0.000520	0.999480	0.985365	98,536	51
5	0.000487	0.999513	0.984852	98,485	48
6	0.000464	0.999536	0.984373	98,437	45
7	0.000453	0.999547	0.983916	98,392	45
8	0.000441	0.999559	0.983470	98,347	43
9	0.000430	0.999570	0.983037	98,304	43
10	0.000419	0.999581	0.982614	98,261	41
11	0.000419	0.999581	0.982202	98,220	41
12	0.000396	0.999604	0.981791	98,179	39
13	0.000427	0.999573	0.981402	98,140	42
14	0.000460	0.999540	0.980983	98,098	45
15	0.000495	0.999505	0.980531	98,053	48
16	0.000533	0.999467	0.980046	98,005	53
17	0.000575	0.999425	0.979524	97,952	56
18	0.000619	0.999381	0.978961	97,896	61
19	0.000667	0.999333	0.978355	97,835	65
20	0.000718	0.999282	0.977702	97,770	70
21	0.000773	0.999227	0.977000	97,700	76
22	0.000833	0.999167	0.976245	97,624	81
23	0.000897	0.999103	0.975432	97,543	87
24	0.000966	0.999034	0.974557	97,456	94
25	0.001041	0.998959	0.973615	97,362	102
26	0.001121	0.998879	0.972602	97,260	109
27	0.001207	0.998793	0.971511	97,151	117
28	0.001300	0.998700	0.970339	97,034	126
29	0.001400	0.998600	0.969077	96,908	136
30	0.001508	0.998492	0.967721	96,772	146
31	0.001624	0.998376	0.966261	96,626	157
32	0.001749	0.998251	0.964692	96,469	169
33	0.001884	0.998116	0.963005	96,300	181
34	0.002029	0.997971	0.961191	96,119	195
35	0.002186	0.997814	0.959240	95,924	210
36	0.002354	0.997646	0.957143	95,714	225
37	0.002535	0.997465	0.954890	95,489	242
38	0.002730	0.997270	0.952470	95,247	260
39	0.002940	0.997060	0.949869	94,987	279
40	0.003166	0.996834	0.947077	94,708	300
41	0.003410	0.996590	0.944078	94,408	322
42	0.003672	0.996328	0.940859	94,086	346
43	0.003954	0.996046	0.937404	93,740	370

*Elementos del Cálculo Actuarial para Vida*

44	0.004258	0.995742	0.933698	93,370	398
45	0.004585	0.995415	0.929722	92,972	426
46	0.004938	0.995062	0.925459	92,546	457
47	0.005317	0.994683	0.920889	92,089	490
48	0.005725	0.994275	0.915993	91,599	524
49	0.006164	0.993836	0.910749	91,075	561
50	0.006637	0.993363	0.905135	90,514	601
51	0.007145	0.992855	0.899128	89,913	643
52	0.007693	0.992307	0.892703	89,270	686
53	0.008282	0.991718	0.885836	88,584	734
54	0.008915	0.991085	0.878499	87,850	783
55	0.009597	0.990403	0.870667	87,067	836
56	0.010330	0.989670	0.862312	86,231	891
57	0.011119	0.988881	0.853404	85,340	948
58	0.011967	0.988033	0.843915	84,392	1,010
59	0.012879	0.987121	0.833816	83,382	1,074
60	0.013860	0.986140	0.823077	82,308	1,141
61	0.014914	0.985086	0.811669	81,167	1,211
62	0.016048	0.983952	0.799564	79,956	1,283
63	0.017265	0.982735	0.786733	78,673	1,358
64	0.018574	0.981426	0.773150	77,315	1,436
65	0.019980	0.980020	0.758789	75,879	1,516
66	0.021490	0.978510	0.743629	74,363	1,598
67	0.023111	0.976889	0.727648	72,765	1,682
68	0.024851	0.975149	0.710831	71,083	1,766
69	0.026720	0.973280	0.693167	69,317	1,852
70	0.028724	0.971276	0.674645	67,465	1,938
71	0.030874	0.969126	0.655267	65,527	2,023
72	0.033180	0.966820	0.635036	63,504	2,107
73	0.035651	0.964349	0.613965	61,397	2,189
74	0.038300	0.961700	0.592077	59,208	2,268
75	0.041136	0.958864	0.569400	56,940	2,342
76	0.044174	0.955826	0.545978	54,598	2,412
77	0.047424	0.952576	0.521860	52,186	2,475
78	0.050902	0.949098	0.497111	49,711	2,530
79	0.054619	0.945381	0.471807	47,181	2,577
80	0.058592	0.941408	0.446037	44,604	2,614
81	0.062834	0.937166	0.419903	41,990	2,638
82	0.067362	0.932638	0.393519	39,352	2,651
83	0.072190	0.927810	0.367011	36,701	2,649
84	0.077337	0.922663	0.340516	34,052	2,634
85	0.082817	0.917183	0.314182	31,418	2,602
86	0.088649	0.911351	0.288162	28,816	2,554
87	0.094850	0.905150	0.262617	26,262	2,491
88	0.101436	0.898564	0.237708	23,771	2,411
89	0.108424	0.891576	0.213595	21,360	2,316
90	0.115832	0.884168	0.190437	19,044	2,206

91	0.123677	0.876323	0.168378	16,838	2,083
92	0.131973	0.868027	0.147553	14,755	1,947
93	0.140737	0.859263	0.128080	12,808	1,803
94	0.149983	0.850017	0.110055	11,005	1,650
95	0.159723	0.840277	0.093548	9,355	1,494
96	0.169970	0.830030	0.078607	7,861	1,336
97	0.180733	0.819267	0.065246	6,525	1,180
98	0.192020	0.807980	0.053454	5,345	1,026
99	0.203837	0.796163	0.043190	4,319	880
100	1.000000	0.000000	0.034386	3,439	3,439

Probabilidades de muerte experiencia mexicana 91 - 98.



Entre las edades 0 y 11 años las probabilidades de muerte decrecen monótonamente hasta alcanzar su mínimo entre los 11 y 12 años. Posteriormente tales probabilidades crecen conforme aumenta la edad hasta llegar a la extinción del grupo, que en este caso a los 101 años de edad.

La probabilidad que tiene un recién nacido de morir antes de cumplir un año de edad ( $q_0$ ), se conoce como Tasa de Mortalidad Infantil (TMI). Es un indicador social muy importante que recibe la atención de muchos estudios y trabajos.

La importancia se debe a que está íntimamente ligada con las condiciones que ofrece una sociedad para el desarrollo de nuevos individuos. Si se tienen de

tasas de mortalidad infantil muy elevadas puede ser consecuencia de problemas económicos y/o de salud que deben ser atendidos inmediatamente.

### **2.1.3 Cálculo de los Elementos de la Tabla de Mortalidad.**

En esta sección se define el cálculo de los elementos básicos de la tabla de mortalidad a partir de los cuales, se pueden derivar los demás elementos del cálculo actuarial para vida. Entonces si el lector solamente memoriza las igualdades de esta sección podrá además deducir las equivalencias de la siguiente sin tener que memorizarlas.

La herramienta de la cual podemos partir para calcular todos los elementos de la tabla de mortalidad es  ${}_x p_0$ .

#### **Cálculo de ${}_t p_x$ .**

${}_t p_x$  = Probabilidad [ (x) llegue con vida a edad  $x+t$  ]

Sean los eventos:

$A_x$  = {Una persona de edad 0 llega con vida a edad  $x$ }

$A_{x+t}$  = {Una persona de edad 0 llega con vida a edad  $x+t$ }

Y por la definición de probabilidad condicional,

$$\begin{aligned} {}_t p_x &= \text{Probabilidad } [ A_{x+t} / A_x ] \\ &= \frac{\text{Probabilidad } [ A_{x+t} \cap A_x ]}{\text{Probabilidad } [ A_x ]} \\ &= \frac{\text{Probabilidad } [ A_{x+t} ]}{\text{Probabilidad } [ A_x ]} \\ &= \frac{{}_{x+t} p_0}{{}_x p_0} \qquad \dots (1) \end{aligned}$$

### Cálculo de ${}_tq_x$ .

${}_tq_x$  = Probabilidad [(x) muera entre las edades  $x$  y  $x+t$ ]

Sean los eventos:

$A_{x,x+t}$  = {(x) llega con vida a edad  $x+t$ }

$B_{x,x+t}$  = {(x) muere entre las edades  $x$  y  $x+t$ }

${}_tq_x$  = Probabilidad [ $B_{x,x+t}$ ]

Como  $(A_{x,x+t})^c = B_{x,x+t}$

${}_tq_x$  = Probabilidad [ $(A_{x,x+t})^c$ ]

Usando el complemento,

$$= 1 - \text{Probabilidad } [A_{x,x+t}]$$

$$= 1 - {}_tP_x \quad \dots (2)$$

### Cálculo de ${}_{n|m}q_x$ .

Para poder determinar  ${}_{n|m}q_x$  utilizando solamente los elementos ya conocidos, es necesario apoyarnos del siguiente teorema de probabilidad:

**Teorema.** Sean A y B eventos del espacio muestral  $\Omega$  se tiene que:

$$\text{Probabilidad } [A \cup B] = \text{Probabilidad } [A] + \text{Probabilidad } [A^c \cap B].$$

Si definimos los eventos:

$A_{x,x+n}$  = {(x) llega con vida a edad  $x+n$ }

$B_{x,x+n}$  = {(x) muere entre las edades  $x$  y  $x+n$ }

$B_{x,x+n|x+n+m}$  = {(x) muere entre las edades  $x+n$  y  $x+n+m$ }

Tenemos que:

$${}_{n+m}q_x = \text{Probabilidad } [(x) muera entre las edades  $x$  y  $x+n+m$ ]$$

$$= \text{Probabilidad } [B_{x,x+n} \cup B_{x,x+n|x+n+m}]$$

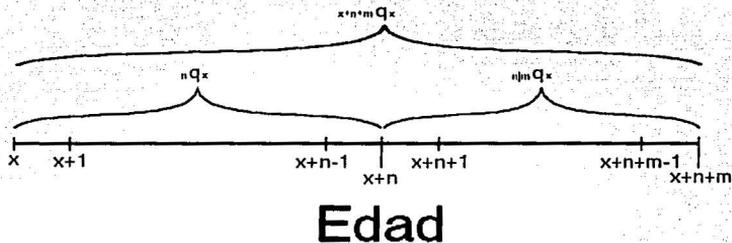
Por el teorema:

$$\begin{aligned}
 &= \text{Probabilidad } [B_{x,x+n}] + \text{Probabilidad } [(B_{x,x+n})^c \cap B_{x,x+n|x+n+m}] \\
 &= \text{Probabilidad } [B_{x,x+n}] + \text{Probabilidad } [A_{x,x+n} \cap B_{x,x+n|x+n+m}] \\
 &= nq_x + n|m q_x
 \end{aligned}$$

Despejando  $n|m q_x$  finalmente tenemos:

$$n|m q_x = n+m q_x - nq_x \quad \dots (3)$$

Este resultado se comprende y se memoriza mejor si se utiliza el siguiente esquema:



### Cálculo de $l_x$ .

Para poder estimar a  $l_x$  es necesario definir provisionalmente una variable aleatoria que cuente el número de sobrevivientes a edad  $x$  y posteriormente determinar la distribución de esta variable aleatoria.

Se define a  $\mathcal{L}_{(x)}$  por la variable aleatoria que denota el número de sobrevivientes a edad  $x$ , de un grupo inicial de  $l_0$  recién nacidos.

Si se recuerda que  $l_x$  se define como el esperado número de sobrevivientes a edad  $x$  de un grupo inicial de  $l_0$  recién nacidos, tenemos entonces que  $l_x$  es la esperanza matemática de la variable aleatoria  $\mathcal{L}_{(x)}$ :

TESIS CON  
 FALLA DE ORIGEN

$$I_x = E[\mathcal{L}(x)]$$

Ahora es necesario determinar la distribución de  $\mathcal{L}(x)$  para poder determinar su esperanza matemática. La observación si un recién nacido sobrevive o no a exactamente la edad  $x$  es un experimento Bernoulli de éxito o fracaso. En este caso el evento éxito es haber sobrevivido a edad  $x$  y el evento fracaso es no haber sobrevivido a edad  $x$ .

Sabemos que la observación simultanea de 2 o más experimentos de Bernoulli independientes implica una distribución Binomial. Entonces  $\mathcal{L}(x)$  tiene una distribución Binomial sobre  $l_0$  recién nacidos:

$$\mathcal{L}(x) \sim \text{Binomial} (n = l_0, p = {}_x p_0)$$

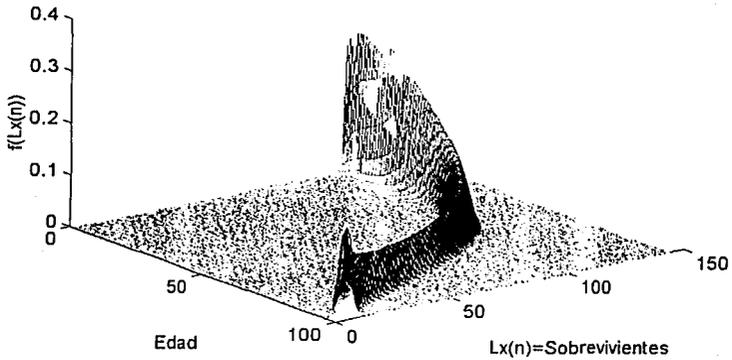
Y como se sabe la esperanza de una Binomial es el producto de sus parámetros,

$$I_x = E[\mathcal{L}(x)] = (l_0) \cdot ({}_x p_0) \quad \dots (4)$$

De esta relación resulta evidente que la distribución  $\mathcal{L}(x)$  denota el número de sobrevivientes a edad  $x$  está directamente determinada por el parámetro  ${}_x p_0$ , el cual no es un valor fijo si no más bien es una función que depende de la edad. Por tal razón se tiene que para cada edad existe una única y diferente distribución de  $\mathcal{L}(x)$ , lo que da como resultado una familia de funciones de distribución que dependen de la edad.

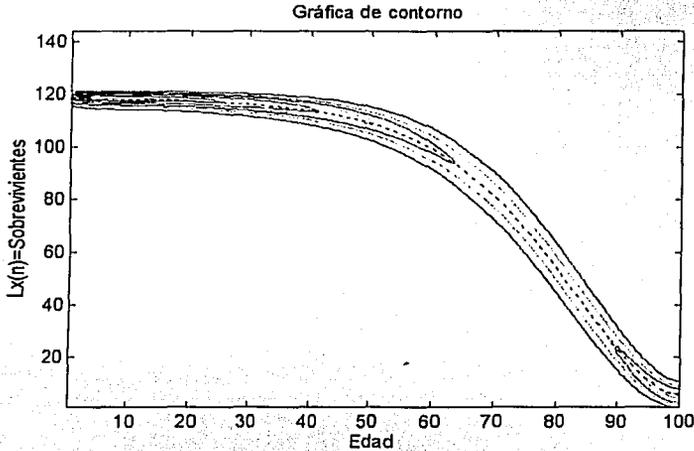
En la siguiente gráfica, creada a partir de un radix igual a 120, se muestra el conjunto de funciones de densidad de  $\mathcal{L}(x)$  para cada una de las edades entre cero y 100. Nótese como cada función de densidad respecto a la edad tiene una forma acampanada, lo que da como resultado una superficie en forma de cordillera montañosa.

Densidades de  $L_x(n)$



Para hacer evidente la relación entre las densidades de  $L_x$  y la función  $l_x$ , en la siguiente gráfica se presentan los cortes horizontales o curvas de nivel de la gráfica anterior. Representada por una curva punteada se encuentra la gráfica de  $l_x$ , la cual pasa exactamente por el centro de las curvas de nivel de las densidades de  $L_x$ .

TESIS CON  
FALLA DE ORIGEN



Como vimos son las probabilidades de supervivencia las que determinan a las  $l_x$  y no como contrariamente se suele pensar. Sin embargo como veremos más adelante existe también el recurso de calcular probabilidades en términos de  $l_x$ .

### Cálculo de ${}_n d_x$ .

Si recordamos la definición,  ${}_n d_x$  es el número esperado de muertes ocurridas entre las edades  $x$  y  $x + n$ , es decir, el número de vivos a edad  $x$  menos el número de vivos a edad  $x+n$ .

$${}_n d_x = l_x - l_{x+n} \quad \dots (5)$$

## 2.1.4 Equivalencias entre los Elementos de la Tabla de Mortalidad.

- Partiendo de (4) podemos encontrar una fórmula recursiva para  $l_x$ :

$$l_{x-n} = l_0 \cdot {}_{x-n} p_0 = l_0 \cdot {}_{x-n} p_0 \cdot 1$$

Si se reemplaza la unidad por  $\frac{xP_0}{xP_0}$ :

$$\Rightarrow l_{x+n} = l_0 \cdot {}_{x+n}p_0 \cdot \frac{xP_0}{xP_0} = l_0 \cdot xP_0 \cdot \frac{x+nP_0}{xP_0} = l_x \cdot {}_n p_x \quad \dots (6)$$

● Despejando  $l_{x+n}$  de (5) y sustituyendo a la edad  $x$  por  $x - n$  podemos fácilmente encontrar otra equivalencia para  $l_x$ :

$$l_x = l_{x-n} \cdot {}_n d_{x-n} \quad \dots (7)$$

● Partiendo también de (5) y utilizando (6) en  $l_{x+n}$  tenemos que:

$${}_n d_x = l_x - l_{x+n} = l_x - l_x \cdot {}_n p_x = l_x (1 - {}_n p_x) = l_x \cdot {}_n q_x \quad \dots (8)$$

● Despejando  ${}_x p_0$  de (4) tenemos:

$${}_x p_0 = \frac{l_x}{l_0}$$

Y sustituyendo  ${}_x p_0$  en (1) encontramos una expresión para  ${}_t p_x$  en términos de  $l_x$ :

$${}_t p_x = \frac{x+tP_0}{xP_0} = \frac{l_{x+t} / l_0}{l_x / l_0} = \frac{l_{x+t} \cdot l_0}{l_x \cdot l_0} = \frac{l_{x+t}}{l_x} \quad \dots (9)$$

● Si en (1) multiplicamos por la unidad sin alterar la identidad tenemos:

$${}_{n+m} p_x = \frac{x+n+mP_0}{xP_0} = \frac{x+n+mP_0}{xP_0} \cdot \frac{x+nP_0}{x+nP_0} = \frac{x+n+mP_0}{x+nP_0} \cdot \frac{x+nP_0}{xP_0} = {}_m p_{x+n} \cdot {}_n p_x \quad \dots (10)$$

● Se puede obtener la siguiente variante aplicando (10) recursivamente:

$${}_t p_x = \prod_{i=0}^{t-1} p_{x+i}, \quad \forall t \in \mathbb{N} \quad \dots (11)$$

Demostración:

$$\prod_{i=0}^{t-1} p_{x+i} = p_x \cdot p_{x+1} \cdots p_{x+t-2} \cdot p_{x+t-1} = \frac{x+1P_0}{xP_0} \cdot \frac{x+2P_0}{x+1P_0} \cdots \frac{x+t-1P_0}{x+t-2P_0} \cdot \frac{x+tP_0}{x+t-1P_0}$$

Reordenando los factores:

$$\Rightarrow \prod_{i=0}^{x-1} P_{x+i} = \frac{1}{x P_0} \cdot \frac{x+1 P_0 \cdot x+2 P_0 \cdots x+1-2 P_0 \cdots x+1-1 P_0}{x+1 P_0 \cdot x+2 P_0 \cdots x+1-2 P_0 \cdots x+1-1 P_0} \cdot \frac{x+1 P_0}{1} = \frac{x+1 P_0}{x P_0} = {}_1 p_x \quad \blacksquare$$

La demostración también es posible aplicando el principio de inducción matemática y la identidad (10).

- La siguiente expresión se obtiene fácilmente despejando a  ${}_1 p_x$  en (2):

$${}_1 p_x = 1 - {}_1 q_x \quad \dots (12)$$

- Del cálculo básico de  ${}_1 q_x$  y aplicando (9) se encuentra la siguiente equivalencia:

$${}_1 q_x = 1 - {}_1 p_x = 1 - \frac{1 - x+1}{1-x} \quad \dots (13)$$

- Si despejamos a  ${}_1 q_x$  de (8) tenemos la siguiente igualdad:

$${}_1 q_x = \frac{{}_1 d_x}{1-x} \quad \dots (14)$$

- Si utilizamos la definición de  ${}_{n|m} q_x$  para descomponer los términos de la siguiente sumatoria, encontraremos un interesante resultado para  ${}_1 q_x$ :

$$\begin{aligned} \sum_{i=0}^t {}_i q_x &= \sum_{i=0}^t ({}_{i+1} q_x - {}_i q_x) = \sum_{i=0}^t (-{}_i q_x + {}_{i+1} q_x) \\ \Rightarrow \sum_{i=0}^t {}_i q_x &= (-{}_0 q_x + {}_1 q_x) + (-{}_1 q_x + {}_2 q_x) + \dots + (-{}_{t-1} q_x + {}_t q_x) + (-{}_t q_x + {}_{t+1} q_x) \end{aligned}$$

Reasociando los términos:

$$\begin{aligned} \Rightarrow \sum_{i=0}^t {}_i q_x &= {}_0 q_x + ({}_1 q_x - {}_1 q_x) + ({}_2 q_x - {}_2 q_x) + \dots + ({}_t q_x - {}_t q_x) + {}_{t+1} q_x \\ \Rightarrow \sum_{i=0}^t {}_i q_x &= 0 + (0) + (0) + \dots + (0) + {}_{t+1} q_x \end{aligned}$$

$$\Rightarrow \sum_{i=0}^t q_x = {}_{t+1}q_x$$

$$\therefore {}_tq_x = \sum_{i=0}^{t-1} q_x \quad \dots (15)$$

● Partiendo de la definición de  ${}_n|mq_x$  y posteriormente utilizando (2) tenemos que:

$${}_n|mq_x = {}_{n+m}q_x - {}_nq_x = (1 - {}_{n+m}p_x) - (1 - {}_np_x) = {}_np_x - {}_{n+m}p_x \quad \dots (16)$$

● El siguiente resultado es muy importante pues relaciona directamente a la probabilidad de muerte diferida con la no diferida. Tal resultado se obtiene si partimos de (16) y posteriormente aplicamos (10) Y (2):

$${}_n|mq_x = {}_np_x - {}_{n+m}p_x = {}_np_x - {}_np_x \cdot {}_m p_{x+n} = {}_np_x (1 - {}_m p_{x+n}) = {}_np_x \cdot q_{x+n} \quad \dots (17)$$

● Si utilizamos a (16) y posteriormente (9) obtenemos a una expresión para  ${}_n|mq_x$  en términos de  $l_x$ .

$${}_n|mq_x = {}_np_x - {}_{n+m}p_x = \frac{l_{x+n}}{l_x} - \frac{l_{x+n+m}}{l_x} = \frac{l_{x+n} - l_{x+n+m}}{l_x} \quad \dots (18)$$

● Otra posible equivalencia para  ${}_n|mq_x$  es inmediata si en (18) se reemplaza a  $l_{x+n} - l_{x+n+m}$  por  ${}_m d_{x+n}$  apoyándose en la definición (5):

$${}_n|mq_x = \frac{{}_m d_{x+n}}{l_x} \quad \dots (19)$$

## 2.2 Función de Mortalidad.

Aunque los elementos de la tabla de mortalidad teóricamente no están restringidos a edades o periodos de tiempo en números enteros, en la práctica la tabla misma se construye sobre la base de periodos anuales enteros.

Pensemos en el problema de calcular la probabilidad que tiene una persona de 25 años de morir antes de tener 30.5 años de edad ( ${}_{5.5}q_{25}$ ) utilizando la tabla de mortalidad "CNSF 2000 - I (1991-1998)". Encontramos que tal probabilidad no se puede calcular de manera exacta y lo más sensato sería aproximar el resultado

mediante la interpolación lineal de las probabilidades  ${}_5q_{25}$  y  ${}_6q_{25}$ . El resultado no sería más que una aproximación, con un error menor o igual a  $|({}_5q_{25} - {}_6q_{25}) / 2|$ .

Se podría pensar erróneamente en solucionar el problema anterior construyendo una tabla de mortalidad con cortes por edad semestrales y no anuales como normalmente se suele hacer. Aunque tal tabla solucionaría de manera exacta nuestro problema específico, no podría solucionar una Infinidad de problemas resultado de trabajar con edades o periodos de tiempo expresados en números reales positivos. Por lo tanto si queremos tener siempre cálculos exactos y no aproximaciones, resulta evidente la necesidad de trabajar con funciones continuas más que con tablas de datos finitos.

Otra razón importante para utilizar una función de mortalidad es que cuando ésta es definida como una variable aleatoria, el método de cálculo de probabilidades, seguros y anualidades contingentes se simplifica y se estandariza ampliamente.

### **2.2.1 Fuerza de Mortalidad $\mu_x$ .**

Las probabilidades de muerte y supervivencia que hemos definido son referidas y tienen sentido solo para intervalos de tiempo, y no así para valores donde no existe ningún intervalo. Si recurrimos a las respectivas definiciones es evidente que  ${}_0p_x = 1$  y  ${}_0q_x = 0$ , con lo cual no se tiene medida alguna de la mortalidad que afecta puntualmente a la edad  $x$ . Es claro entonces que los elementos hasta aquí definidos no son suficientes para el desarrollo de una función continua de mortalidad.

Por lo tanto el primer paso para el desarrollo de una función continua de mortalidad es tener una medida instantánea o puntual de la mortalidad.

El concepto que a continuación se define tiene la feliz característica de proporcionar una medida práctica de la mortalidad puntual además de tener un sentido teórico:

La *fuerza de mortalidad* denotada por  $\mu_x$  se define como la tasa instantánea de mortalidad que afecta a una persona de exactamente edad  $x$ . Como ya explicamos anteriormente una tasa es un promedio unitario respecto al tiempo,

idea con la cual podemos escribir a  $\mu_x$  mediante la siguiente expresión matemática:

$$\mu_x = \lim_{h \rightarrow 0^+} \frac{h q_x}{h} \quad \dots (20)$$

Aunque esta definición pudiera dar la idea de que  $\mu_x$  y  $q_x$  toman valores semejantes la verdad es que no ocurre así. Mientras  $h q_x$  solo puede tomar valores entre 0 y 1,  $\mu_x$  y  $h q_x/h$  cuando  $h$  es menor a 1 pueden tomar cualquier valor real positivo. El siguiente ejemplo ilustra esta situación.

Supongamos que el día 1° de enero del 2000 se registraron 1782 nacimientos en la república mexicana, de los cuales 100 murieron antes del primer día de edad. Entonces se tiene que  $\frac{h q_x}{h} = \frac{100 q_0}{1/365} = (100/1792) \cdot 365 \approx 20.482604$

lo cual es mucho mayor a 1.

La gran importancia de  $\mu_x$  no radica en el hecho de ser una medida puntual de la mortalidad, si no que a partir de ella misma se pueden calcular el 100% de los elementos que se utilizan en el cálculo actuarial para vida. En otras palabras, si es posible obtener o determinar una función para representar a  $\mu_x$  automáticamente se tiene resuelto todo el cálculo actuarial para vida.

## 2.2.2 Función de Supervivencia $s(x)$ .

### Definición y Características.

De igual manera que ocurre entre una probabilidad de muerte  ${}_t q_x$  y una probabilidad de supervivencia  ${}_t p_x$ , las funciones de supervivencia y de mortalidad son complementarias entre sí y cada una se puede calcular en función de la otra.

Por lo tanto el construir una función de mortalidad es equivalente a construir su función espejo de supervivencia.

La función definida por la probabilidad que tiene un recién nacido de llegar con vida a edad  $x$ , denotada por  $s(x)$ , recibe el nombre de *función de supervivencia*.

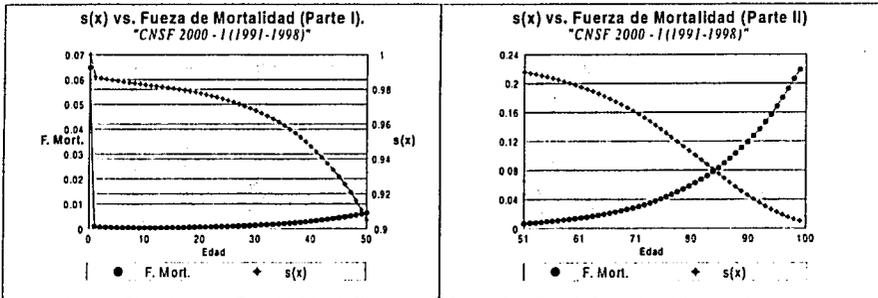
$$s(x) = {}_x p_0. \quad \dots (21)$$

Las propiedades de  $s(x)$  son consecuencia directa de las leyes de probabilidad y la distribución de la mortalidad de una población. Se listan a continuación:

- $s(x): \{\mathbb{R}^+ \cup \{0\}\} \longrightarrow [0,1]$ .
- $s(x)$  es continua y diferenciable en el intervalo abierto  $(0,w)$ .
- $s(0) = {}_0P_0 = 1$ .
- $s(x) = 0 \quad \forall \quad x \geq w$ .
- $s(x)$  es monótona decreciente, es decir  $x_1 \geq x_2 \Leftrightarrow s(x_1) \leq s(x_2)$ .

### Relación entre la Mortalidad y $s(x)$ .

Existe una relación recíproca entre la función de supervivencia y la mortalidad que actúa sobre la población en cada edad. La pendiente de la recta tangente para la función  $s(x)$  a la edad  $x$  es inversamente proporcional a la mortalidad que afecta a la población a esa edad. Esto ocasiona que en edades en las que la mortalidad es muy elevada, la función  $s(x)$  decaiga muy rápidamente como se puede ver en la siguiente gráfica:



Este comportamiento se demuestra y toma sentido en el siguiente resultado, el cual comúnmente se utiliza como una definición alterna para  $\mu_x$  por que expresa a éste en términos de  $s(x)$ . Si aplicamos la definición de  $\mu_x$  y también (2) tenemos que:

TESIS CON  
FALLA DE ORIGEN

$$\mu_x = \lim_{h \rightarrow 0^+} \frac{h q_x}{h} = \lim_{h \rightarrow 0^+} \left( \frac{1 - {}_h p_x}{h} \right)$$

Si ahora aplicamos (1) :

$$\Rightarrow \mu_x = \lim_{h \rightarrow 0^+} \frac{1 - \frac{{}_x p_0}{{}_x p_0} \frac{{}_x p_0}{{}_x p_0}}{\frac{h}{1}} = \lim_{h \rightarrow 0^+} \frac{{}_x p_0 - {}_{x+h} p_0}{\frac{{}_x p_0}{1} \cdot \frac{h}{1}}$$

Por la ley del sándwich y factorizando -1:

$$\Rightarrow \mu_x = \lim_{h \rightarrow 0^+} \frac{{}_x p_0 - {}_{x+h} p_0}{{}_x p_0 \cdot h} = \lim_{h \rightarrow 0^+} \frac{-1 \cdot ({}_{x+h} p_0 - {}_x p_0)}{{}_x p_0 \cdot h}$$

Si ahora sacamos del límite los factores que no dependen de h y sustituimos a  ${}_x p_0$  por  $s(x)$  con base en la definición (21):

$$\Rightarrow \mu_x = \left( \frac{-1}{{}_x p_0} \right) \cdot \lim_{h \rightarrow 0^+} \frac{{}_{x+h} p_0 - {}_x p_0}{h} = \left( \frac{-1}{s(x)} \right) \cdot \lim_{h \rightarrow 0^+} \frac{s(x+h) - s(x)}{h}$$

Finalmente utilizando la definición de derivada sobre el límite encontramos que:

$$\Rightarrow \mu_x = \left( \frac{-1}{s(x)} \right) \cdot s'(x) = \frac{-s'(x)}{s(x)} \quad \dots (22)$$

### Cálculo de $s(x)$ en Términos de $\mu_x$ .

Para calcular a  $s(x)$  en términos de  $\mu_x$  debemos pensar a la expresión (22) como una ecuación diferencial lineal homogénea de primer orden y resolverla. Afortunadamente la resolución de tan sencilla ecuación diferencial no requiere más que conceptos básicos de integración y funciones primitivas. Para que el resultado convenga a nuestro objetivo, debemos añadir a la ecuación el valor inicial  $s(0) = 1$  obteniendo así el siguiente sistema:

$$\begin{cases} s'(x) + \mu_x \cdot s(x) = 0 \\ s(0) = 1 \end{cases}$$

Resolvamos entonces paso a paso este sencillo sistema:

$$\frac{s'(t)}{s(t)} = -\mu_t$$

En general para toda función  $f(x)$  se cumple  $(\ln(f(x)))' = f'(x)/f(x)$

$$\Rightarrow \frac{d}{dt} \ln(s(t)) = -\mu_t$$

Por el Teorema Fundamental del Cálculo segunda parte II<sup>(3)</sup>:

$$\Rightarrow \int_0^x \left( \frac{d}{dt} \ln(s(t)) \right) dt = \int_0^x -\mu_t dt$$

$$\Rightarrow \ln(s(t)) \Big|_0^x = - \int_0^x \mu_t dt$$

Si ahora valuamos la Integral y utilizamos el valor inicial  $s(0) = 1$  para un recién nacido:

$$\Rightarrow \ln(s(x)) - \ln(s(0)) = \ln(s(x)) - \ln(1) = - \int_0^x \mu_t dt$$

Como  $\ln(1) = 0$

$$\Rightarrow \ln(s(x)) = - \int_0^x \mu_t dt$$

Finalmente aplicando la función exponencial a ambos lados de la igualdad:

$$\Rightarrow s(x) = {}_x p_0 = e^{-\int_0^x \mu_t dt} \quad \dots (23)$$

El resultado (23) es muy importante, pues contar con una equivalencia de  ${}_x p_0$  en términos de  $\mu_x$ , permite calcular a los demás elementos de la tabla de mortalidad en términos de  $\mu_x$ .

<sup>(3)</sup> Teorema Fundamental del Cálculo parte II:

Si  $F$  es cualquier antiderivada de  $f$  en el intervalo  $[a, b]$  entonces

$$\int_a^b f(x) dx = F(b) - F(a)$$

### 2.2.3 Función de Mortalidad $F(x)$ .

#### Definición y Características.

La función definida por la probabilidad que tiene un recién nacido de morir antes de tener la edad  $x$ , denotada por  $F(x)$ , recibe el nombre de *función de mortalidad*.

$$F(x) = {}_xq_0. \quad \dots (24)$$

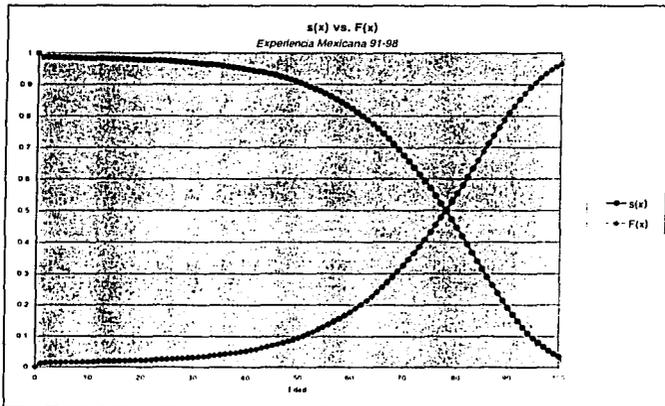
Adoptando la típica nomenclatura entre funciones primitivas y funciones derivadas, se denota a la derivada de  $F(x)$  por  $f(x)$ , es decir

$$F'(x) = f(x) \quad \dots (25)$$

En este momento es donde podemos verificar que  $F(x)$  y  $s(x)$  son realmente funciones complementarias, ya que si en la definición anterior aplicamos la definición (2):

$$F(x) = {}_xq_0 = 1 - {}_xp_0 = 1 - s(x) \quad \dots (26)$$

La relación anterior la podemos verificar al hacer la grafica de  $s(x)$  vs.  $F(x)$ :

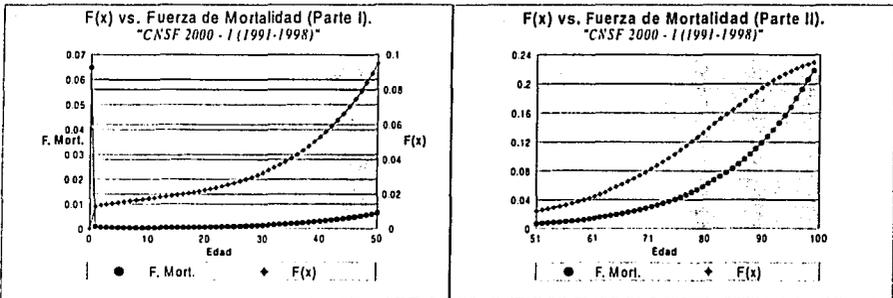


Tal como ocurre con la función  $s(x)$  las propiedades de  $F(x)$  son consecuencia directa de las leyes de probabilidad y la distribución de la mortalidad de una población:

- $F(x): \{\mathbb{R} + \cup \{0\}\} \longrightarrow [0, 1]$ .
- $F(x)$  es continua y diferenciable en el intervalo abierto  $(0, w)$ .
- $F(0) = {}_0q_0 = 0$ .
- $F(x) = 1 \quad \forall x \geq w$ .
- $F(x)$  es monótona creciente, es decir  $x_1 \geq x_2 \Leftrightarrow F(x_1) \geq F(x_2)$ .

### Relación entre la Mortalidad y $F(x)$ .

De la misma manera que ocurre con la función de supervivencia  $s(x)$ , existe una relación entre la función de mortalidad  $F(x)$  y la mortalidad que actúa sobre la población a cada edad. La pendiente de la recta tangente para la función  $F(x)$  en una determinada edad  $x$ , es directamente proporcional a la mortalidad que afecta a la población en esa edad. Esto ocasiona que en edades en las que la mortalidad es muy elevada, la función  $F(x)$  crezca muy rápidamente como se puede ver en la siguiente gráfica:



Esta relación se explica en el siguiente resultado. Si se utiliza (22) y posteriormente se sustituye (26):

$$\mu_x = \frac{-s'(x)}{s(x)} = \frac{-(1-F(x))'}{1-F(x)} = \frac{F'(x)}{1-F(x)} = \frac{f(x)}{1-F(x)} \quad \dots (27)$$

### **Cálculo de F(x) en Términos de $\mu_x$ .**

El cálculo de F(x) a partir de  $\mu_x$  resulta muy sencillo si recordamos que  $F(x) + s(x) = 1$ :

$$F(x) = 1 - s(x) = 1 - e^{-\int_0^x \mu_t dt} \quad \dots (28)$$

## **2.2.4 Cálculo de los Elementos de la Tabla de Mortalidad con $\mu_x$ .**

Como ya mencionamos anteriormente,  $\mu_x$  es una poderosa herramienta con la cual podemos calcular todos los elementos de la tabla de mortalidad. Esta labor es más sencilla por el hecho que ahora contamos con una equivalencia entre  ${}_x p_0$  y  $\mu_x$ .

### **Cálculo de ${}_t p_x$ .**

Si partimos de la definición de  ${}_t p_x$  (1) y aplicamos (23) tenemos:

$${}_t p_x = \frac{{}_{x+t} P_0}{{}_x P_0} = \frac{e^{-\int_0^{x+t} \mu_s ds}}{e^{-\int_0^x \mu_s ds}}$$

Utilizando leyes de los exponentes:

$$\Rightarrow {}_t p_x = e^{-\int_0^{x+t} \mu_s ds} \cdot e^{\int_0^x \mu_s ds} = e^{-\int_0^{x+t} \mu_s ds + \int_0^x \mu_s ds} = e^{-\left(\int_0^{x+t} \mu_s ds - \int_0^x \mu_s ds\right)}$$

Como sabemos que  $\int_A^B = -\int_B^A$ :

$$\Rightarrow {}_t p_x = e^{-\left(-\int_0^x \mu_s ds + \int_0^{x+t} \mu_s ds\right)} = e^{-\left(\int_x^{x+t} \mu_s ds + \int_0^x \mu_s ds\right)}$$

Como sabemos que  $\int_A^B + \int_B^C = \int_A^C$ :

$$\Rightarrow {}_tP_x = e^{-\int_t^{x+t} \mu_{x+s} ds}$$

Esta igualdad se puede simplificar aun más aplicando el cambio de variable  $u = s - x$ . Teorema de cambio de variable:

$$\text{Si } u = g(s) \Rightarrow \int_a^b f(g(s)) \cdot g'(s) = \int_{g(a)}^{g(b)} f(u) \cdot du$$

$$\begin{aligned} \text{Con } u &= s - x, & \Rightarrow du &= ds \\ s &= x, & \Rightarrow u &= (x) - x = 0 \\ s &= x + t, & \Rightarrow u &= (x+t) - x = t \end{aligned}$$

$$\Rightarrow {}_tP_x = e^{-\int_t^{x+t} \mu_{x+s} ds}$$

$$\Rightarrow {}_tP_x = e^{-\int_0^t \mu_{x+u} du} = e^{-\int_0^t \mu_{x+u} dt} \quad \dots (29)$$

### Cálculo de ${}_tq_x$ .

Si partimos de la definición de  ${}_tq_x$  y aplicamos (29) tenemos que:

$${}_tq_x = 1 - {}_tP_x = 1 - e^{-\int_0^t \mu_{x+u} du} \quad \dots (30)$$

### Cálculo de ${}_{n|m}q_x$ .

De la identidad (16) y aplicando posteriormente (29) tenemos que:

$${}_{n|m}q_x = {}_n P_x - {}_{n+m} P_x = e^{-\int_0^n \mu_{x+t} dt} - e^{-\int_0^{n+m} \mu_{x+t} dt} \quad \dots (31)$$

### Cálculo de $l_x$ .

Con la definición (4) y aplicando de nuevo (23):

$$l_x = l_{0:x} p_0 = l_0 \cdot e^{-\int_0^x \mu_t dt} \quad \dots (32)$$

### **Cálculo de ${}_n d_x$ .**

Si se utiliza la definición (5) y posteriormente (23) se tiene la siguiente expresión para  ${}_n d_x$ :

$${}_n d_x = l_x - l_{x+n} = l_0({}_x p_0 - {}_{x+n} p_0) = l_0 \left( e^{-\int_0^x \mu_t ds} - e^{-\int_0^{x+n} \mu_t ds} \right) \quad \dots (33)$$

## 2.3 Variables Aleatorias para Mortalidad.

En esta sección se estudiarán variables aleatorias específicamente definidas para representar la distribución de la mortalidad de una población.

Se podría tener la impresión de que los elementos actuariales hasta aquí vistos forman las bases suficientes para el desarrollo del cálculo actuarial para vida, y que definir una variable aleatoria esta de sobra. Sin embargo, existen algunos elementos inherentes del cálculo actuarial cuya apropiada definición y explicación solo es posible utilizando la adecuada variable aleatoria, como es por ejemplo la esperanza completa de vida a edad  $x$  ( $e_x$ ).

Por otra parte la esperanza matemática de una variable aleatoria resulta ser la mejor herramienta para el cálculo del costo de seguros de vida y anualidades contingentes. Además el manejo de los elementos actuariales ya definidos se simplifica y se estandariza ampliamente mediante la implementación de variables aleatorias.

### 2.3.1 Variable Aleatoria $X$ ~ Edad de Muerte.

La variable aleatoria que representa la edad a la que muere un recién nacido se denota por  $X$  ~ Edad de muerte. Aunque sus parámetros no se incluyen en la notación, se sugiere que podrían ser los siguientes:

- Población o poblaciones a las cuales representa.  
Por ejemplo México.
- Período de tiempo de observación sobre el cual se construyeron los datos.  
Por ejemplo 1991-1998
- Restricción o restricciones del conjunto de la población que se tomó para observación.  
Por ejemplo solo hombres, no fumadores, sin restricción, etc.

Los posibles valores para  $X$  son el cero y cualquier número real positivo. Por esta razón podemos afirmar que  $X$  es una variable aleatoria continua.

### **Función de distribución.**

Utilizando la definición de función de distribución para una variable aleatoria, la función de distribución para  $X \sim$  Edad de muerte está dada por la probabilidad de que  $X$  tome cualquier valor menor o igual a la edad  $x$ . Rescribiendo esta definición como a continuación se muestra, se puede encontrar una adecuada definición para  $F_x(x)$  en términos de funciones ya conocidas:

$$F_x(x) = \text{Probabilidad } (X \leq x)$$

$$\Rightarrow F_x(x) = \text{Probabilidad } (X < x) + \text{Probabilidad } (X = x)$$

$$\Rightarrow F_x(x) = \text{Probabilidad (Un recién nacido muera entre las edades 0 y } x) \\ + \text{Probabilidad (Un recién nacido muera exactamente a la edad } x)$$

Como para toda variable aleatoria continua Probabilidad  $(X=x) = 0$ .

$$\Rightarrow F_x(x) = {}_xq_0 + 0 = F(x) \quad \dots (34)$$

Por lo tanto la función de distribución de la variable aleatoria  $X$  es justamente la función de mortalidad  $F(x)$ .

### **Función de densidad.**

Gracias a la definición, sabemos que la función de densidad para una variable aleatoria continua se puede obtener derivando a la función de distribución. Como en este caso resulta que la función de distribución para  $X \sim$  Edad de muerte es la función de mortalidad  $F(x)$ , tenemos que:

$$f_x(x) = F'(x) = f(x) \quad \dots (35)$$

Con esta idea y por (27) se puede encontrar la función de densidad para  $X$ :

$$\mu_x = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{s(x)}$$

$$\Rightarrow f(x) = \mu_x \cdot s(x) = \mu_x \cdot {}_xP_0 \quad \dots (36)$$

Y ahora si se aplica (23) para reemplazar  $s(x)$ :

$$\Rightarrow f_x(x) = f(x) = \mu_x e^{-\int_0^x \mu_t dt} \quad \dots (37)$$

### 2.3.2 Variable Aleatoria $T_{(x)} \sim$ Tiempo Restante de Vida.

La variable aleatoria que representa el tiempo restante de vida que tiene una persona de edad  $x$  se denota por  $T_{(x)} \sim$  Tiempo restante de vida. Los parámetros que caracterizan esta variable aleatoria son los mismos que los de  $X \sim$  Edad de muerte, pero tiene un parámetro adicional que es la edad actual de individuo  $x$ .

La principal diferencia con  $X \sim$  Edad de muerte, es que  $T_{(x)}$  no representa la edad en la que muere un individuo si no más bien el tiempo que le queda de vida a un individuo de edad  $x$  conocida. Entonces para  $T_{(x)}$  la edad  $x$  no es un valor sino más bien un parámetro de la variable aleatoria.

$T_{(x)}$  también es continua y puede tomar cualquier valor mayor o igual a cero. Además  $X \sim$  Edad de muerte es un caso particular de  $T_{(x)}$ , puesto que  $T_{(0)} = X$ .

#### Función de distribución.

La función de distribución de  $T_{(x)}$  se denota actuarialmente por  $G(t)$ , es decir:

$$F_{T_{(x)}}(t) = G(t)$$

Para encontrar la distribución de  $T_{(x)}$  se puede usar el mismo procedimiento para  $X \sim$  Edad de muerte, utilizando la definición de función de distribución para una variable aleatoria. En este caso la edad  $x$  es un parámetro de la variable aleatoria y  $t$  representa el tiempo en años que le queda de vida a la persona de edad  $x$ :

$$G(x) = \text{Probabilidad } (T_{(x)} \leq t)$$

$$\Rightarrow G(x) = \text{Probabilidad } (T_{(x)} < t) + \text{Probabilidad } (T_{(x)} = t)$$

Por ser  $T_{(x)}$  una variable aleatoria continua:

$$\Rightarrow G(x) = \text{Probabilidad } (T_{(x)} < t)$$

$\Rightarrow G(x) = \text{Probabilidad (Una persona de edad } x \text{ le queden menos de } t \text{ años de vida)}$

$\Rightarrow G(x) = \text{Probabilidad (Una persona de edad } x \text{ muera antes de tener la edad } x + t)$

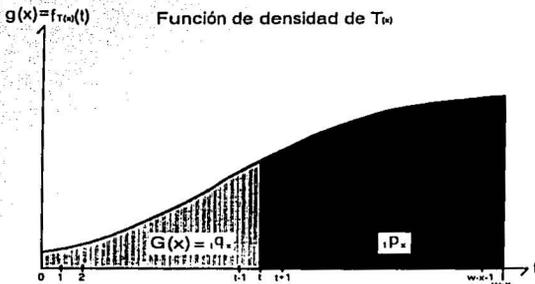
$$\Rightarrow G(x) = {}_1q_x \quad \dots (38)$$

### Función de densidad.

A la función de densidad de la variable aleatoria  $T_{(x)}$  actuarialmente se le denota por  $g(t)$ , es decir:

$$f_{T_{(x)}}(t) = g(t)$$

Las equivalencias (38),  $G(w-x) = 1$  y  ${}_t p_x + {}_t q_x = 1$ , inducen una relación entre la función de distribución, la función de densidad y las probabilidades  ${}_t p_x$  y  ${}_t q_x$  tal y como se muestra en la siguiente figura:



Nuevamente por ser  $T_{(x)}$  continua, el método para encontrar la función de densidad consiste en derivar a la función de distribución. Un método poco común para derivar es utilizando la definición sobre el límite, pero en este caso se justifica su uso para poder encontrar una identidad adicional para la función de densidad:

$$f_{T_{(x)}}(t) = g(t) = \frac{d({}_t q_x)}{dt} = \lim_{h \rightarrow 0} \frac{{}_{t+h} q_x - {}_t q_x}{h} = \lim_{h \rightarrow 0^+} \frac{{}_{t+h} q_x - {}_t q_x}{h}$$

Por la definición de  ${}_n | m q_x$  (3) tenemos que:

$$\Rightarrow g(t) = \lim_{h \rightarrow 0^+} \frac{{}_t | h q_x}{h} \quad \dots (39)$$

TESIS CON  
FOLIO DE ORIGEN

Si aplicamos la identidad (17) se puede simplificar:

$$\Rightarrow g(t) = \lim_{h \rightarrow 0^+} \frac{{}_t p_x \cdot h q_{x+t}}{h} = {}_t p_x \cdot \lim_{h \rightarrow 0^+} \frac{h q_{x+t}}{h}$$

Ahora si se utiliza la definición de la fuerza de mortalidad (20):

$$\Rightarrow g(t) = {}_t p_x \cdot \mu_{x+t} \quad \dots (40)$$

Y con (29) podemos reemplazar a  ${}_t p_x$  por  $\mu_x$ :

$$\Rightarrow g(t) = e^{-\int_0^t \mu_{x+s} ds} \cdot \mu_{x+t} \quad \dots (41)$$

En este punto es donde se pueden ver concretamente las ventajas de  $T_{(x)}$ , pues su función de densidad  $g(t)$  permite desarrollar todos los elementos de la tabla de mortalidad.

### **2.3.3 Variable Aleatoria $K_{(x)} \sim$ Años Enteros Restantes de Vida.**

Una forma rápida de comprender a la variable aleatoria  $K(x) \sim$  Años enteros restantes de vida es verla como la aproximación discreta de la variable aleatoria  $T(x) \sim$  Tiempo restante de vida.

Se define por  $K_{(x)} \sim$  Años enteros restantes de vida a la variable aleatoria que denota los años enteros que le restan de vida a una persona de edad  $x$ . Los parámetros que caracterizan esta variable aleatoria son exactamente los mismos que los de la variable aleatoria  $T_{(x)}$  los cuales ya fueron descritos.

De la misma manera que ocurre para la variable aleatoria  $T_{(x)}$ , la edad  $x$  no es un valor sino más bien un parámetro de la variable aleatoria  $K_{(x)}$ .

$K_{(x)}$  es una variable aleatoria discreta y solo puede tomar valores entre los números enteros no negativos. Es común que el parámetro  $x$  sea truncado a un número entero, pero esto en realidad no es una regla; la edad  $x$  puede tomar el valor de cualquier número entero positivo o cero.

## Función de densidad.

Para las variables aleatorias discretas es mejor en la mayoría de los casos determinar primero la función de densidad y posteriormente la función de distribución. La principal razón es que a diferencia de las variables aleatorias continuas, la función de densidad para las variables aleatorias discretas tiene una interpretación real que suele ser muy sencilla de determinar.

En este caso a la función de densidad de la variable aleatoria  $K_{(x)}$  no se le asigna ninguna notación en particular, por lo que solo utilizaremos la notación típica que se le suele dar a las funciones de densidad.

El método para determinar la función de densidad para variables aleatorias discretas es aplicando directamente la definición:

$$f_{K_{(x)}}(t) = \text{Probabilidad}(K_{(x)} = t)$$

$$\Rightarrow f_{K_{(x)}}(t) = \text{Probabilidad}(K_{(x)} = t)$$

$\Rightarrow f_{K_{(x)}}(t) = \text{Probabilidad}(\text{Que a una persona de edad } x \text{ le restan exactamente } t \text{ años enteros de vida})$

$\Rightarrow f_{K_{(x)}}(t) = \text{Probabilidad}(\text{Una persona de edad } x \text{ sobreviva hasta la edad } x + t \text{ y además muera antes de tener la edad } x + t + 1)$

$$\Rightarrow f_{K_{(x)}}(t) = {}_t|q_x \quad \text{con } t \in \mathbb{N} \cup \{0\} \quad \dots (42)$$

La aproximación de  $T_{(x)}$  por  $K_{(x)}$  se refleja incluso en sus funciones de densidad, pues como podemos observar  $f_{K_{(x)}}(t) = {}_t|q_x$  y  $f_{T_{(x)}}(t) = \lim_{h \rightarrow 0^+} \frac{{}_t|q_x}{h}$

## Función de distribución.

La función de distribución de una variable aleatoria discreta en el valor  $t$ , es la sumatoria de las densidades para los valores menores o iguales a  $t$ :

$$F_{K_{(x)}}(t) = \text{Probabilidad}(K_{(x)} \leq t)$$

$$\Rightarrow F_{K_{(x)}}(t) = f_{K_{(x)}}(0) + f_{K_{(x)}}(1) + \dots + f_{K_{(x)}}(t)$$

Si recordamos que  $f_{K(x)}(t) = {}_t|q_x$  (42):

$$\Rightarrow F_{K(x)}(t) = {}_0|q_x + {}_1|q_x + \dots + {}_t|q_x$$

Finalmente con la identidad (15) se llega a que:

$$\Rightarrow F_{K(x)}(t) = {}_{t+1}q_x \quad \text{con } t \in \mathbb{N} \cup \{0\} \quad \dots (43)$$

Nótese como las distribuciones de  $K(x)$  y de  $T(x)$  son casi idénticas.

## 2.4 Otras Funciones Vitales.

Existe un grupo de funciones elementales cuya definición depende de las variables aleatorias  $T(x)$  y  $K(x)$ . Como se podrá observar y se mencionó anteriormente,  $T(x)$  es la base para la mayoría de los cálculos. Es por esto que  $T(x)$  es una función que nunca debemos perder de vista.

### 2.4.1 Esperanza Completa de Vida.

Se define a la esperanza completa de vida a edad  $x$ , denotada por  ${}^o e_x$ , como la esperanza matemática de la variable aleatoria  $T(x)$ :

$${}^o e_x = E[T(x)] \quad \dots (44)$$

En términos más claros lo que ésta definición quiere decir, es que la esperanza completa de vida es el número de años que en promedio vivirán las personas que tienen edad  $x$ .

En contraste con el cálculo actuarial donde la esperanza completa de vida casi no se utiliza, en demografía es un elemento muy importante para comparar las condiciones de vida entre dos poblaciones o para analizar el desarrollo de una población específica a través del tiempo.

Como es costumbre, para determinar  ${}^o e_x$  se puede hacer aplicando directamente la definición de esperanza matemática para la variable aleatoria  $T(x)$ . Recordando que la densidad de  $T(x)$  en  $t$  ( $f_{T(x)}(t)$ ) se denota por  $g(t)$  y que solo es distinta de cero para el intervalo  $[0, w-x)$ :

$${}^{\circ}e_x = E[T_{(x)}] = \int_0^{\infty} t \cdot f_{T_{(x)}}(t) dt = \int_0^{w-x} t \cdot g(t) dt$$

Si se aplica la fórmula de integración por partes  $\int v du = v \cdot u - \int u \cdot dv$  con  $v=t$ ,  $dv=dt$ ,  $du=g(t)$  y  $u=G(t)$ ,

$$\Rightarrow {}^{\circ}e_x = t \cdot G(t) \Big|_0^{w-x} - \int_0^{w-x} G(t) dt$$

$$\Rightarrow {}^{\circ}e_x = [(w-x) \cdot G(w-x) - 0 \cdot G(0)] - \int_0^{w-x} G(t) dt$$

Puesto que  $G(t) = {}_tq_x$  no es difícil ver que  $G(w-x) = {}_{w-x}q_x = 1$ ,

$$\Rightarrow {}^{\circ}e_x = [(w-x) \cdot 1 - 0] - \int_0^{w-x} {}_tq_x dt = (w-x) - \int_0^{w-x} (1 - {}_tp_x) dt$$

Y como la integral de la suma es la suma de las integrales

$$\Rightarrow {}^{\circ}e_x = (w-x) - \left( \int_0^{w-x} 1 dt - \int_0^{w-x} {}_tp_x dt \right) = (w-x) - \left( (w-x) - \int_0^{w-x} {}_tp_x dt \right)$$

$$\Rightarrow {}^{\circ}e_x = \int_0^{w-x} {}_tp_x dt \quad \dots (45)$$

## 2.4.2 Tiempo Medio de Vida $m(x)$ .

Se define al Tiempo medio de vida  $m(x)$  como el cuantil 0.5 ( $\xi_{0.5}$ ) o mediana de la variable aleatoria  $T_{(x)}$ . En términos de probabilidades lo que esta definición quiere decir, es que  $m(x)$  es el punto en el tiempo en el que una persona de edad  $x$  tiene la misma probabilidad de haber muerto que de haber sobrevivido.

$$m(x) = \{t \mid \text{Probabilidad } (T_{(x)} \leq t) = 1/2\} = \{t \mid G(t) = 1/2\} \quad \dots (46)$$

$$\Rightarrow m(x)q_x = 1/2$$

$$\Rightarrow m(x)p_x = 1/2$$

Para determinar a  $m(x)$  no existe un método general. Para cada caso dependiendo de las características de la situación como puede ser  $\mu_x$ , se puede escoger entre despear a  $m(x)$ , o aproximarlo.

### 2.4.3 Años Persona Vividos ${}_tL_x$ .

Se denota a  ${}_tL_x$  por el número total de años vividos entre las edades  $x$  y  $x + t$ . La expresión algebraica equivalente a la definición se cita a continuación:

$${}_tL_x = \int_x^{x+t} l_s ds = \int_0^t l_{x+s} ds \quad \dots (47)$$

Si utilizamos (4) y la definición de  $s(x)$  se encuentra otra posible equivalencia para  ${}_tL_x$  es la siguiente:

$${}_tL_x = \int_0^t \int_0^{x+s} p_0 ds = l_0 \cdot \int_0^t s(x+s) ds \quad \dots (48)$$

Cuando  $t = 1$  se omite escribirla y la notación se reduce a  $L_x$ . Puesto que la integral de una función en un período unitario es el valor promedio de la misma,  $L_x$  también es la media de los valores que toma la función  $l_y$  en el intervalo  $x$  y  $x + 1$ .

Por esta razón, el promedio exacto  $L_x$  se suele aproximar mediante el promedio aritmético  $\frac{1}{2}(l_x + l_{x+1})$ :

$$L_x = \frac{l_x + l_{x+1}}{2} \quad \dots (49)$$

### 2.4.4 Total Años Persona Vividos $T_x$ .

Se define a  $T_x$  como el número total de años vividos desde la edad  $x$  hasta la muerte, por el grupo de  $l_x$  personas que sobrevivieron a edad  $x$ . Como se puede ver,  $T_x$  es un caso particular de  ${}_tL_x$  con  $t = w$ :

$$T_x = {}_wL_x = \int_x^{w-x} l_t dt \quad \dots (50)$$

Nótese que el cociente de  $T_x$  entre  $l_x$  es justamente la esperanza completa de vida:

$$\frac{T_x}{l_x} = \frac{\int_0^{w-x} l_{x+t} dt}{l_x} = \int_0^{w-x} \frac{l_{x+t}}{l_x} dt$$

Sabemos que  ${}_t p_x = \frac{l_{x+t}}{l_x}$  y por la definición de  $e_x^{\circ}$ ,

$$\Rightarrow \frac{T_x}{l_x} = \int_0^{w-x} {}_t p_x dt = e_x^{\circ} \quad \dots (51)$$

## 2.4.5 Tasa Central de Mortalidad $m_x$ .

Se define a la tasa central de mortalidad  $m_x$  por el resultado del cociente de  $d_x$  entre  $L_x$ .

$$m_x = \frac{d_x}{L_x} \quad \dots (52)$$

Es conocida por tasa central de mortalidad, pues es el cociente del número de muertos dentro de un periodo de un año, con el promedio de  $l_x$  dentro del mismo periodo.

## 2.5 Tablas Selectas de Mortalidad.

Existen algunos seguros de vida los cuales para su venta como prerequisite requieren la aplicación de un examen médico para los solicitantes. A los solicitantes que acreditan satisfactoriamente el examen y que no requieren condiciones especiales de asegurabilidad se les conoce como *vidas selectas*.

Para las vidas selectas, la edad actual del individuo se denota por  $[x] + t$ , donde  $[x]$  representa la edad en la que se realizó el proceso de selección y  $t$  representa el número de años que han pasado desde la selección.

Es natural pensar que a consecuencia de la selección, las distribuciones de muerte para una vida selecta sean diferentes que para una vida ordinaria. De hecho una vida selecta tiene mejores condiciones de supervivencia, cualidad que se reduce paulatinamente conforme la edad se aleja del momento en que se efectuó la selección:

$$q_{[x]} < q_{[x-1]+1} < q_{[x-2]+2} \dots$$

Al periodo de tiempo dentro del cual las probabilidades de muerte tienen una desviación significativa a causa del proceso de selección se le conoce como *periodo de selección*. Generalmente al periodo de selección se le denota por la letra  $r$ , y se define formalmente mediante la siguiente expresión:

$$r = \{ \min. \{K\} \mid |q_{[x]+k} - q_{[x-j]+k+j}| < \varepsilon \quad \forall x \text{ donde } [x] \text{ esté definido}, \\ \text{con } k \in \mathbb{N}, 0 < j < x \} \quad \dots (53)$$

Resulta evidente que la distribución de la mortalidad para una vida selecta no depende solamente de la edad del individuo si no además del tiempo que ha transcurrido desde la selección. Por tal razón para el manejo de vidas selectas es necesario contar con funciones de 2 variables independientes o con arreglos de datos de doble entrada como son las tablas selectas de mortalidad.

Una *tabla selecta* de mortalidad es el ordenamiento de  $r$  tablas de mortalidad, en el que cada una contiene las probabilidades de muerte para cada año del periodo de selección. Una tabla selecta y última de mortalidad es un conjunto de  $r + 1$  tablas de mortalidad conformado por la unión de una tabla selecta de mortalidad y la tabla última de mortalidad.

La *tabla última* de mortalidad es aquella que contiene las probabilidades de muerte para vidas que fueron selectas pero que ya no están dentro del periodo de selección. Por lo general se toma como tabla última a la tabla para edad alcanzada (la tabla ordinaria para vidas no selectas) aunque no necesariamente.

Por convención, en las columnas de la tabla selecta se han dispuesto las probabilidades que comparten un mismo tiempo desde la selección, y en las filas se han dispuesto las probabilidades que comparten la misma edad de selección. De esta manera la tabla selecta y última de mortalidad debe contener  $r + 1$  columnas donde cada una es en sí una tabla de mortalidad, y hasta un máximo de  $w - r - 1$  filas.

El seguimiento de una vida selecta dentro de la tabla selecta siempre se hace horizontalmente de izquierda a derecha dentro de los primeros  $r$  años y después verticalmente de arriba abajo.

	Selecta					Última
$q_{[x+r]}$	0	1	$t$	$r-1$	$r$	
$[x]$	$q_{[x]}$	$q_{[x]+1}$	$q_{[x]+t}$	$q_{[x]+r-1}$	$q_{x+r}$	
$[x+1]$	$q_{[x+1]}$	$q_{[x+1]+1}$	$q_{[x+1]+t}$	$q_{[x+1]+r-1}$	$q_{x+1+r}$	
$[x+s]$	$q_{[x+s]}$	$q_{[x+s]+1}$	$q_{[x+s]+t}$	$q_{[x+s]+r-1}$	$q_{x+s+r}$	
$[w-r-1]$	$q_{[w-r-1]}$	$q_{[w-r-1]+1}$	$q_{[w-r-1]+t}$	$q_{[w-r-1]+r-1}$	$q_{w-1} = 1$	

➔ Seguimiento de una vida selecta a través de la tabla selecta y última.

### 2.5.1 Probabilidades Selectas.

Como consecuencia de la selección tenemos que  ${}_t p_{[x]} \neq \frac{{}_{x+t}P_0}{{}_x P_0}$ . Por tal razón muchas de las equivalencias ya vistas no son válidas para vidas selectas, con lo que es necesario redefinir herramientas de trabajo que nos permitan manejar adecuadamente las vidas selectas.

Las equivalencias que a continuación se presentan permiten el cálculo de los elementos actuariales para vidas selectas, a partir de tablas selectas que estén escritas en términos de  $q_{[x]}$ ,  $p_{[x]}$  o  $l_{[x]}$ . El desarrollo o la demostración de tales equivalencias se omitirá por ser análogos, aunque más complicados, que los que aparecen en las secciones anteriores.

No es recomendable utilizar tablas en formato  $l_{[x]}$ , ni calcular probabilidades selectas en términos de  $l_{[x]}$  puesto que se pierde la referencia del radix  $l_0$ . Sin embargo puede hacerse un "no recomendado" desarrollo análogo si se considera a  $l_{[x]}$  como un radix selecto.

TESIS CON  
FALLA DE ORIGEN

**Cálculo de  ${}_tP_{[x]+k}$ .**

$${}_tP_{[x]+k} = \frac{{}_{k+t}P_{[x]}}{{}_kP_{[x]}} \quad \dots (54)$$

$${}_tP_{[x]+k} = \prod_{i=0}^{t-1} p_{[x]+i}, \text{ cont } t, k \in \mathbb{N} \quad \dots (55)$$

$${}_tP_{[x]+k} = \frac{l_{[x]+k+t}}{l_{[x]+k}} \quad \dots (56)$$

$${}_tP_{[x]+k} = 1 - {}_tq_{[x]+k} \quad \dots (57)$$

**Cálculo de  ${}_tq_{[x]+k}$ .**

$${}_tq_{[x]+k} = 1 - {}_tP_{[x]+k} \quad \dots (58)$$

$${}_tq_{[x]+k} = \sum_{i=0}^{t-1} {}_i q_{[x]+k}, \text{ cont } t, k \in \mathbb{N} \quad \dots (59)$$

**Cálculo de  ${}_{n|m}q_{[x]+k}$ .**

$${}_{n|m}q_{[x]+k} = {}_{n+m}q_{[x]+k} - {}_nq_{[x]+k} \quad \dots (60)$$

$${}_{n|m}q_{[x]+k} = {}_n p_{[x]+k} - {}_{n+m} p_{[x]+k} \quad \dots (61)$$

$${}_{n|m}q_{[x]+k} = \frac{l_{[x]+k+n} - l_{[x]+k+n+m}}{l_{[x]+k}} \quad \dots (62)$$

$${}_{n|m}q_{[x]+k} = {}_n p_{[x]+k} \cdot {}_m q_{[x]+k+n} \quad \dots (64)$$

**Cálculo de  $l_{[x]+k}$ .**

Si se define a  $l_{[x]}$  como un *radix selecto* se puede determinar  $l_{[x]+k}$  de la siguiente manera:

$$l_{[x]+k} = l_{[x]} \cdot {}_k p_{[x]} \quad \dots (65)$$

Si partimos del hecho que para poder construir la tabla selecta y última en términos de  $l_{[x]}$  se debe cumplir que  $l_{[x]+r+k} = l_{[x+k]+r}$ , donde  $r$  es el periodo de selección y  $k \geq 0$ :

$$\Rightarrow l_{[x]} \cdot {}_{k+r}p_{[x]} = l_{[x+k]} \cdot r p_{[x+k]}$$

$$\Rightarrow l_{[x+k]} = \frac{l_{[x]} \cdot {}_{k+r}p_{[x]}}{r p_{[x+k]}} \quad \dots (66)$$

**Cálculo de  ${}_n d_{[x]+k}$**

$${}_n d_{[x]+k} = l_{[x]+k} - l_{[x]+k+n} \quad \dots (67)$$

**Nota: Debe recordarse que:**

$$p_{[x]+r+k} \approx p_{x+r+k},$$

$$q_{[x]+r+k} \approx q_{x+r+k},$$

$$l_{[x]+r+k} \approx l_{x+r+k},$$

donde  $r$  es el periodo de selección y  $k \geq 0$ .

**2.5.2 Ejemplos de Cálculo de Probabilidades Selectas:**

Tabla Selecta Ejemplo. (*)		Periodo selección $r = 4$			
$[x]$	$q_{[x]}$	$q_{[x]+1}$	$q_{[x]+2}$	$q_{[x]+3}$	$q_{x+4}$
20	0.0005026	0.0006648	0.0007497	0.0008432	0.000966
21	0.0005411	0.0007164	0.0008073	0.000908	0.001041
22	0.0005831	0.0007714	0.0008694	0.0009785	0.001121
23	0.0006279	0.0008308	0.0009369	0.0010537	0.001207
24	0.0006762	0.0008953	0.0010089	0.0011346	0.0013
25	0.0007287	0.0009641	0.0010863	0.001222	0.0014
26	0.0007847	0.001038	0.00117	0.001316	0.001508

(\*) Valores calculados con la tabla de mortalidad "CNSF 2000 - I (1991-1998)" y los factores para vidas selectas hombres que aparecen en el trabajo "Tablas de Mortalidad México 2000" realizado por la AMIS.

Con la Tabla Selecta Ejemplo:

- 1) Calcular  ${}_3p_{[20]}$

- 2) Calcular  ${}_3P_{|20|+3}$
- 3) Calcular  ${}_3P_{|20|+3}$
- 4) Calcular  ${}_3P_{|23|+3}$
- 5) Calcular  ${}_3q_{|23|+3}$
- 6) Calcular  ${}_3|q_{|20|+3}$
- 7) Calcular  ${}_3|_3q_{|20|+3}$
- 8) Escribir la Tabla Selecta Ejemplo en términos de  $l_{[x]}$ , con un radix selecto de 100,000.

Resultados:

- 1)  ${}_3P_{|20|} = P_{|20|} \cdot P_{|20|+1} \cdot P_{|20|+2} = (1-q_{|20|}) \cdot (1-q_{|20|+1}) \cdot (1-q_{|20|+2})$   
 $\Rightarrow {}_3P_{|20|} = (1-0.0005026) \cdot (1-0.00066478) \cdot (1-0.0007497) = 0.99808413$
- 2)  ${}_3P_{|20|+3} = 1 - q_{|20|+3} = (1-0.00084318) = 0.99915682$
- 3)  ${}_3P_{|20|+3} = ({}_3+{}_3P_{|20|}) / ({}_3P_{|20|}) = ({}_6P_{|20|}) / ({}_3P_{|20|})$   
 $\Rightarrow {}_3P_{|20|+3} = (P_{|20|} \cdot P_{|20|+1} \cdot P_{|20|+2} \cdot P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5}) / (P_{|20|} \cdot P_{|20|+1} \cdot P_{|20|+2})$   
 $\Rightarrow {}_3P_{|20|+3} = P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5} \approx P_{|20|+3} \cdot P_{20+4} \cdot P_{20+5}$   
 $\Rightarrow {}_3P_{|20|+3} \approx (1-0.00084318) \cdot (1-0.000966) \cdot (1-0.001041) = 0.99715252$
- 4)  ${}_3P_{|23|+3} = ({}_6P_{|23|}) / ({}_3P_{|23|}) = P_{|23|+3} \cdot P_{|23|+4} \cdot P_{|23|+5} \approx P_{|23|+3} \cdot P_{23+4} \cdot P_{23+5}$   
 $\Rightarrow {}_3P_{|23|+3} \approx (1-0.00105374) \cdot (1-0.001207) \cdot (1-0.0013)$   
 $\Rightarrow {}_3P_{|23|+3} \approx 0.99644347$
- 5)  ${}_3q_{|23|+3} = 1 - {}_3P_{|23|+3} = 1 - 0.99644347 = .00355653$
- 6)  ${}_3|q_{|20|+3} = {}_3P_{|20|+3} - {}_3+1P_{|20|+3} = {}_3P_{|20|+3} - {}_4P_{|20|+3}$   
 $\Rightarrow {}_3|q_{|20|+3} = (P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5}) - (P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5} \cdot P_{|20|+6})$   
 $\Rightarrow {}_3|q_{|20|+3} = (P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5}) \cdot (1 - P_{|20|+6})$   
 $\Rightarrow {}_3|q_{|20|+3} \approx (P_{|20|+3} \cdot P_{20+4} \cdot P_{20+5}) \cdot (1 - P_{20+6})$   
 $\Rightarrow {}_3|q_{|20|+3} \approx (1-0.00084318) \cdot (1-0.000966) \cdot (1-0.001041) \cdot (0.001121)$   
 $\Rightarrow {}_3|q_{|20|+3} \approx 0.99707266$
- 7)  ${}_3|_3q_{|20|+3} = {}_3P_{|20|+3} - {}_3+3P_{|20|+3}$   
 $\Rightarrow {}_3|_3q_{|20|+3} = (P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5}) -$   
 $\quad - (P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5} \cdot P_{|20|+6} \cdot P_{|20|+7} \cdot P_{|20|+8})$   
 $\Rightarrow {}_3|_3q_{|20|+3} = (P_{|20|+3} \cdot P_{|20|+4} \cdot P_{|20|+5}) \cdot (1 - P_{|20|+6} \cdot P_{|20|+7} \cdot P_{|20|+8})$   
 $\Rightarrow {}_3|_3q_{|20|+3} \approx (P_{|20|+3} \cdot P_{20+4} \cdot P_{20+5}) \cdot (1 - P_{20+6} \cdot P_{20+7} \cdot P_{20+8})$   
 $\Rightarrow {}_3|_3q_{|20|+3} \approx (P_{|20|+3} \cdot P_{20+4} \cdot P_{20+5}) \cdot (1 - P_{20+6} \cdot P_{20+7} \cdot P_{20+8})$

$$\begin{aligned} \Rightarrow {}_3|3q_{[20]+3} &\approx (1-0.00084318) \cdot (1-0.000966) \cdot (1-0.001041) \cdot \\ &\quad \cdot [1-(1-0.001121) \cdot (1-0.001207) \cdot (1-0.0013)] \\ \Rightarrow {}_3|3q_{[20]+3} &\approx 0.0036133041 \end{aligned}$$

8) Si  $l_{[20]} = 100,000$

$$\begin{aligned} \Rightarrow l_{[21]} &= 100,000 \cdot ({}_5p_{[20]} / {}_4p_{[21]}) = 100,000 \cdot (0.995242101 / 0.995992547) \\ &= 99925 \\ \Rightarrow l_{[22]} &= 100,000 \cdot ({}_6p_{[20]} / {}_4p_{[22]}) = 100,000 \cdot (0.994126435 / 0.995683927) \\ &= 99844 \\ \Rightarrow l_{[23]} &= 100,000 \cdot ({}_7p_{[20]} / {}_4p_{[23]}) = 100,000 \cdot (0.992926524 / 0.995352268) \\ &= 99756 \\ \Rightarrow l_{[24]} &= 100,000 \cdot ({}_8p_{[20]} / {}_4p_{[24]}) = 100,000 \cdot (0.99163572 / 0.994994998) \\ &= 99662 \\ \Rightarrow l_{[25]} &= 100,000 \cdot ({}_9p_{[20]} / {}_4p_{[25]}) = 100,000 \cdot (0.99024743 / 0.994610467) \\ &= 99561 \\ \Rightarrow l_{[26]} &= 100,000 \cdot ({}_{10}p_{[20]} / {}_4p_{[26]}) = 100,000 \cdot (0.988754137 / 0.994196648) \\ &= 99453 \end{aligned}$$

Así podemos continuar hasta llenar todos los valores

Tabla Selecta Ejemplo.		Periodo selección = 4			
[x]	$l_{[x]}$	$l_{[x]+1}$	$l_{[x]+2}$	$l_{[x]+3}$	$l_{x+4}$
20	100000	99950	99883	99808	99724
21	99925	99871	99799	99719	99628
22	99844	99786	99709	99622	99525
23	99756	99693	99611	99517	99412
24	99662	99595	99505	99405	99292
25	99561	99488	99393	99285	99163
26	99453	99375	99272	99156	99025

### 2.5.3 Propuesta de Función Selecta de supervivencia.

#### Definición.

Si es posible contar con una tabla de mortalidad selecta, es mejor aún poder trabajar con una función de dos variables independientes que me permita el adecuado manejo de vidas selectas.

Se define a la *función selecta de supervivencia* por la función continua dada por:

$$s([x],k) = {}_kP_{[x]} \quad \dots (68)$$

**Cálculo de los elementos actuariales para vidas selectas con  $s([x], k)$ .**

$${}_tP_{[x]:k} = \frac{s([x], k+t)}{s([x], k)}$$

$${}_tq_{[x]:k} = 1 - \frac{s([x], k+t)}{s([x], k)}$$

$${}_{n|m}q_{[x]:k} = \frac{s([x], k+n) - s([x], k+n+m)}{s([x], k)}$$

$$l_{[x]:t} = l_{[x]} \cdot s([x], t)$$

$${}_td_{[x]:k} = l_{[x]} \cdot (s([x], k) - s([x], k+t))$$

En este momento contamos con los elementos suficientes para avanzar a la siguiente parte del texto donde revisaremos los métodos de ajuste.

### III. Método de Ajuste AMIS-AMA.

El objetivo de los dos siguientes capítulos es el presentar y analizar métodos de ajuste para tablas de mortalidad, así como comparar ventajas y desventajas de cada uno de éstos. En este capítulo se expondrá un método utilizado por la AMIS (Asociación Mexicana de Instituciones de Seguros, A.C.) y la AMA (Asociación Mexicana de Actuarios A.C.), el cual también se encuentra documentado en la publicación "Tablas de Mortalidad México 2000" que realizaron ambas instituciones.

Como ya se explicó anteriormente, pero vale la pena recordar por ser tema central del presente capítulo, el ajustar una tabla de mortalidad consiste en encontrar una expresión matemática la cual represente adecuadamente la distribución de la mortalidad expresada por la tabla misma.

Es recomendable que todo método de ajuste se base en el desarrollo de expresiones matemáticas directamente sobre la fuerza de mortalidad, ya que a partir de ésta es posible encontrar de manera inmediata expresiones matemáticas para todos los elementos del cálculo actuarial para Seguro de Personas.

#### 3.1 Primera Ley de Makeham.

Dentro del conjunto de hipótesis de leyes matemáticas que rigen el comportamiento de la mortalidad, el más importante es el que se conoce como Ley de Makeham. Esta ley, atribuye la mortalidad que afecta a una población a dos principales causas o componentes. El primero es un componente azaroso que afecta por igual a todas las edades (A), y el segundo es un componente exponencial en función de la edad (BC<sup>x</sup>):

$$\mu_x = A + BC^x$$

La ley de Makeham, que data desde 1860, es uno de los supuestos más poderosos para el cálculo actuarial para los Seguros de Personas por ajustar con gran fidelidad y muy pocos parámetros la mortalidad de una población. Estas dos propiedades hacen que esta Ley no sea solo un instrumento teórico sino además una herramienta práctica que se utiliza cotidianamente. Un claro ejemplo que ilustra esta situación se puede apreciar en cálculo para vidas conjuntas, donde el supuesto de la Ley de Makeham hace posible manejar probabilidades de supervivencia para todo un grupo de personas con solo una persona de edad equivalente.

### 3.2 Cálculo de ${}_t p_x$ Utilizando la Ley de Makeham.

Si aplicamos de la ley de Makeham  $\mu_x = A + BC^x$  en la definición  ${}_t p_x = e^{-\int_0^t \mu_{x+s} ds}$  podemos encontrar una expresión matemática para  ${}_t p_x$  bajo el siguiente procedimiento:

$${}_t p_x = e^{-\int_0^t A + BC^{x+s} ds}$$

Puesto que una función primitiva de  $f(s) = C^s$  es  $F(s) = C^s / \ln(C)$  se tiene:

$$\Rightarrow {}_t p_x = e^{-\left[As + \frac{BC^{x+s}}{\ln(C)}\right]_0^t} = e^{-\left[A(t-0) + \frac{BC^t}{\ln(C)}(C^t - C^0)\right]} = e^{-At - \frac{BC^t}{\ln(C)}(C^t - 1)}$$

Para simplificar la expresión anterior, se suele aplicar la sustitución de las siguientes constantes:  $\ln(s) = -A$  y  $\ln(g) = -\frac{B}{\ln(C)}$ .

$$\Rightarrow {}_t p_x = e^{\ln(s) + \ln(g) C^t \cdot (C^t - 1)} = e^{\ln(s^t)} e^{\ln(g^{C^t(C^t-1)})}$$

$$\Rightarrow {}_t p_x = s^t \cdot g^{C^t \cdot (C^t - 1)} \quad \dots (69)$$

### 3.3 Determinando los Parámetros $s$ , $g$ , y $C$ .

El presente método se basa en la resolución de un sistema de tres ecuaciones basadas en (69) para determinar los tres parámetros  $s$ ,  $g$  y  $C$ .

El primer paso consiste en dividir el intervalo de edades ( $[x, x+n]$ ) para las cuales se va a ajustar la curva en 3 subintervalos iguales. Posteriormente para cada uno de los subintervalos se deben estimar las probabilidades de supervivencia  ${}_1 p_x$ ,  ${}_2 p_x$ ,  ${}_3 p_x$  y  ${}_n p_x$ , utilizando los datos de la tabla de mortalidad. Para este propósito nos podemos apoyar en el resultado (11) del capítulo 2.

Si asumimos que se cumple la ley de Makeham, entonces podemos aplicar a las probabilidades arriba estimadas la igualdad (69) obteniendo como resultado un sistema de 3 ecuaciones. La sustitución  $t = (1/3) \cdot n$  simplifica al sistema mencionado de la siguiente manera:

$$\begin{cases} {}_t p_x = s^t g^{C^x(C^t-1)} \\ {}_t p_{x+t} = s^t g^{C^{x+t}(C^t-1)} \\ {}_t p_{x+2t} = s^t g^{C^{x+2t}(C^t-1)} \end{cases}$$

A cada una de las ecuaciones del sistema anterior le aplicamos la función continua y biyectiva  $f(x)=\ln(x)$ , con lo que se logra separar los factores en sumandos y los coeficientes en factores. Esto abre el camino para posteriormente reducir el número de incógnitas y el número de ecuaciones del sistema.

$$\begin{cases} \ln({}_t p_x) = \ln(s^t) + \ln(g^{C^x(C^t-1)}) \\ \ln({}_t p_{x+t}) = \ln(s^t) + \ln(g^{C^{x+t}(C^t-1)}) \\ \ln({}_t p_{x+2t}) = \ln(s^t) + \ln(g^{C^{x+2t}(C^t-1)}) \end{cases}$$

Por propiedades de  $\ln(x)$ ,

$$\Rightarrow \begin{cases} \ln({}_t p_x) = t \cdot \ln(s) + C^x (C^t - 1) \cdot \ln(g) \\ \ln({}_t p_{x+t}) = t \cdot \ln(s) + C^{x+t} (C^t - 1) \cdot \ln(g) \\ \ln({}_t p_{x+2t}) = t \cdot \ln(s) + C^{x+2t} (C^t - 1) \cdot \ln(g) \end{cases} \quad \dots (3.1)$$

De este último sistema, se resta la primera ecuación a la segunda ecuación y análogamente la segunda ecuación a la tercera ecuación. Como resultado el sistema se reduce a uno de 2 ecuaciones en el que ya no aparece el término  $t \cdot \ln(s)$ :

$$\begin{cases} \ln({}_t p_{x+t}) - \ln({}_t p_x) = C^{x+t} (C^t - 1) \cdot \ln(g) - C^x (C^t - 1) \cdot \ln(g) \\ \ln({}_t p_{x+2t}) - \ln({}_t p_{x+t}) = C^{x+2t} (C^t - 1) \cdot \ln(g) - C^{x+t} (C^t - 1) \cdot \ln(g) \end{cases}$$

Factorizando los términos semejantes y aplicando la propiedad del logaritmo en la que  $\ln(A) - \ln(B) = \ln(A/B)$  tenemos que:

$$\begin{cases} \ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right) = (C^t - 1) \cdot \ln(g) \cdot [C^{x+t} - C^x] \\ \ln\left(\frac{{}_tP_{x+2t}}{{}_tP_{x+t}}\right) = (C^t - 1) \cdot \ln(g) \cdot [C^{x+2t} - C^{x+t}] \end{cases}$$

Ahora si se utiliza la ley de los exponentes en la que  $A^{B+C} = A^B \cdot A^C$  se tiene que:

$$\begin{cases} \ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right) = (C^t - 1) \cdot \ln(g) \cdot C^x [C^t - 1] \\ \ln\left(\frac{{}_tP_{x+2t}}{{}_tP_{x+t}}\right) = (C^t - 1) \cdot \ln(g) \cdot C^{x+t} [C^t - 1] \end{cases}$$

$$\Rightarrow \begin{cases} \ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right) = (C^t - 1)^2 \cdot \ln(g) \cdot C^x \\ \ln\left(\frac{{}_tP_{x+2t}}{{}_tP_{x+t}}\right) = (C^t - 1)^2 \cdot \ln(g) \cdot C^{x+t} \end{cases} \quad \dots (3.ii)$$

Para eliminar una de las dos ecuaciones del sistema y determinar el coeficiente C, dividimos la segunda ecuación entre la primera ecuación:

$$\Rightarrow \frac{\ln\left(\frac{{}_tP_{x+2t}}{{}_tP_{x+t}}\right)}{\ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right)} = \frac{(C^t - 1)^2 \cdot \ln(g) \cdot C^{x+t}}{(C^t - 1)^2 \cdot \ln(g) \cdot C^x} = C^t$$

Finalmente podemos despejar C para lo cual se debe aplicar  $\sqrt{\quad}$  a ambos lados de la ecuación. Nótese que para que sea válido este despeje,  $\ln\left(\frac{{}_tP_{x+2t}}{{}_tP_{x+t}}\right)$  y  $\ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right)$  deben ser al mismo tiempo positivos o negativos, de lo contrario podría resultar en raíces imaginarias.

$$\therefore C = \sqrt{\frac{\ln\left(\frac{{}_tP_{x+2t}}{{}_tP_{x+t}}\right)}{\ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right)}} \quad \dots (70)$$

Para encontrar el parámetro  $g$ , basta con despejarlo de cualquiera de las ecuaciones del sistema (3.ii). Es importante cerciorarse de haber encontrado previamente el valor de  $C$  por que este es necesario para determinar a  $g$ . Si despejamos a  $g$  de la primera ecuación del sistema (3.ii) nos queda:

$$g = e^{\frac{\ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right)}{C^t \cdot (C^t - 1)}} \quad \dots (71)$$

De igual manera, para determinar  $s$  es necesario despejarla de cualquiera de las ecuaciones del sistema (3.i), habiendo obtenido previamente los valores de  $C$  y  $g$ . Si despejamos a  $s$  de la primera ecuación del sistema nos queda que:

$$s = e^{\frac{\ln({}_tP_x) - C^t \cdot (C^t - 1) \cdot \ln(g)}{t}} \quad \dots (72)$$

En caso de que se deseen encontrar los parámetros originales  $A$  y  $B$  de la ley de Makeham, es posible determinarlos mediante las igualdades  $A = -\ln(s)$  y  $B = -\ln(C) \cdot \ln(g)$ .

Es necesario aclarar que mediante un razonamiento paralelo, es posible ajustar también la tabla de mortalidad utilizando la Ley de Gompertz " $\mu_x = BC^x$ ". Para tal propósito es necesario partir de un sistema de dos ecuaciones y simplificar de manera similar.

Además de las leyes de Gompertz y Makeham existen otras Leyes al comportamiento de  $\mu_x$ : el modelo de Moivre  $\mu_x = (w - x)^{-1}$  y el modelo de Weibull  $\mu_x = kx^n$ .

### 3.4 Aplicación.

Como ejemplo se tomó la Tabla "CNSF 2000 - I (1991-1998)" en el intervalo de edades entre 13 y 99 años.

Con una edad inicial  $x = 13$  y periodos de  $t = 29$  años obtenemos las tres ecuaciones iniciales del sistema:

$${}_tP_x = {}_{29}P_{13} = \prod_{k=0}^{28} P_{13+k} = 0.958688889$$

$${}_tP_{x+t} = {}_{29}P_{42} = \prod_{k=0}^{28} p_{42+k} = 0.696455684$$

$${}_tP_{x+2t} = {}_{29}P_{71} = \prod_{k=0}^{28} p_{71+k} = 0.05247624$$

Se tiene que:

$$C = \sqrt[29]{\frac{\ln\left(\frac{{}_{29}P_{71}}{{}_{29}P_{42}}\right)}{\ln\left(\frac{{}_{29}P_{42}}{{}_{29}P_{13}}\right)}} = \sqrt[29]{\frac{\ln\left(\frac{0.05247624}{0.696455684}\right)}{\ln\left(\frac{0.696455684}{0.958688889}\right)}} = 1.074758259$$

$$g = e^{\frac{\ln\left(\frac{{}_tP_{x+t}}{{}_tP_x}\right)}{C^t \cdot (C^t - 1)^2}} = e^{\frac{\ln\left(\frac{0.696455684}{0.958688889}\right)}{1.074758259^{13} \cdot (1.074758259^{29} - 1)^2}} = 0.997513803$$

$$s = e^{\frac{\ln({}_tP_x) - C^t(C^t - 1) \cdot \ln(g)}{t}} = e^{\frac{\ln(0.958688889) - 1.074758259^{13}(1.074758259^{29} - 1) \cdot \ln(0.997513803)}{29}} = 1.000099177$$

$$A = -\ln(s) = -\ln(1.000099177) = -9.9171761097 \times 10^{-5}$$

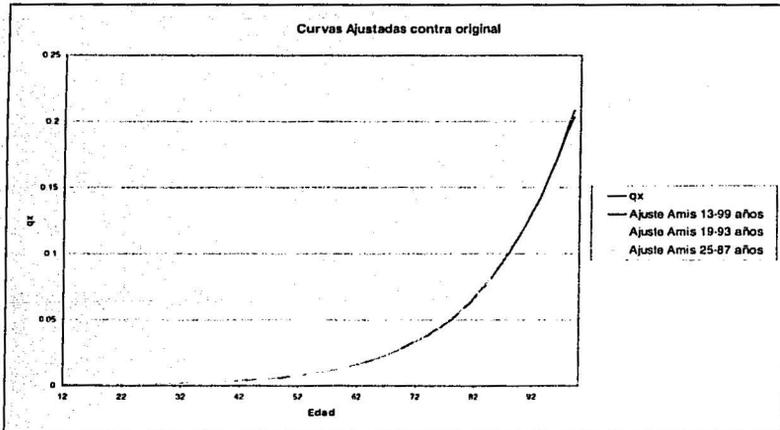
$$B = -\ln(C) \cdot \ln(g) = -\ln(1.074758259) \cdot \ln(0.997513803) = 1.7946745633 \times 10^{-4}$$

### 3.5 Resultados.

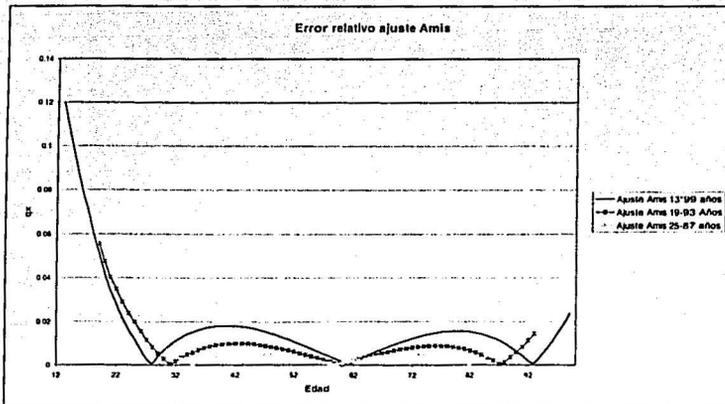
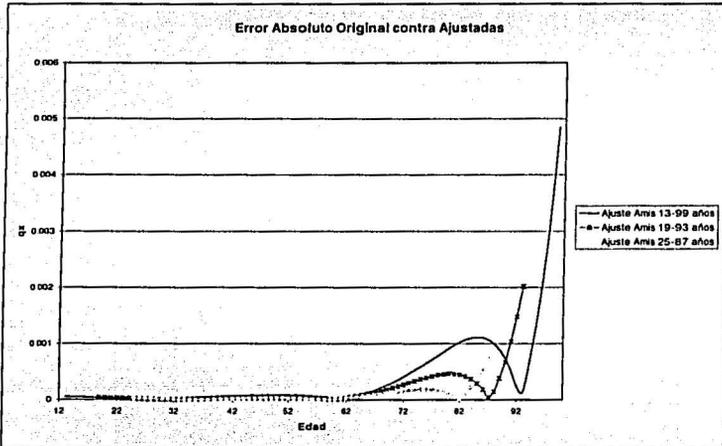
Para efecto de futuras comparaciones, se aplicó el método para tres diferentes rangos de edades de la Tabla de Mortalidad "CNSF 2000 - I (1991-1998)". Se obtuvo la siguiente tabla de parámetros:

Periodo de Edades.	$\mu_x = A + BC^x$			${}_1p_x = s^t \cdot g^{C^t(C^t-1)}$	
	A	B	C	s	g
13-99 años.	$-9.9171761097 \cdot 10^{-05}$	$1.7946745633 \cdot 10^{-04}$	1.074758259	1.000099177	0.997513803
19-93 años	$-8.539077002 \cdot 10^{-05}$	$1.7366780489 \cdot 10^{-04}$	1.075300295	1.000085394	0.997610739
25-87 años	$-7.440812981 \cdot 10^{-05}$	$1.6964435205 \cdot 10^{-04}$	1.075694857	1.000074411	0.997677749

En la figura que se encuentra a continuación, se muestran las gráficas de las tasas de mortalidad de la tabla Experiencia Mexicana 91-98 y de las tasas calculadas con los parámetros antes obtenidos:



El error entre las tasas calculadas y las tasas de la tabla de mortalidad se encuentra representado en las siguientes dos gráficas. Se debe observar particularmente como el error crece rápidamente conforme se amplía el rango de edades a ajustar:



### **3.6 Alcances y Limitaciones.**

#### **Alcances.**

Es un método que en la práctica resulta sencillo y rápido de usar, ya que se reduce a la sola aplicación de las fórmulas tal y como se hizo en el ejemplo del presente capítulo.

Mediante éste método, no se requiere del uso de computadoras o paquetes especiales de computación; una calculadora científica es más que suficiente.

La función resultante es muy confiable pero a la vez muy sencilla, sobre todo cuando no se necesitan ajustar rangos de edades muy extensos.

El hecho de contar con una representación de la mortalidad en términos de la Ley de Makeham, permite sacar provecho de múltiples análisis y simplificaciones que se han realizado sobre la Ley de Makeham.

#### **Limitaciones.**

Como se observó anteriormente la Ley de Makeham y por consecuencia también este método, presentan desviaciones importantes cuando el rango de edades a ajustar es muy amplio. Por ésta razón, es común encontrar que se utilice la Ley de Makeham para graduar tablas de mortalidad sólo para rangos de edades no muy amplios.

El no suponer y manejar un error entre el modelo y los datos originales, tal y como ocurre por ejemplo en regresión lineal, ocasiona que cuando los datos no se comportan fielmente conforme a la Ley de Makeham el modelo resultado no represente a la tabla para varias edades.

Por basarse éste método en la Ley de Makeham, con parámetros fijos y preestablecidos, no permite ingresar nuevas variables al modelo para mejorarlo y por consecuencia mejorar el ajuste. Por lo tanto el modelo no es maleable y no hay manera de agregar variables a la fórmula para reducir el error.

El método no permite ajustar tablas cuyas tasas de mortalidad crezcan para algunos rangos de edades y decrezcan para otros. Es decir, toda tabla con un comportamiento errático no podrá ser ajustada.

## IV. Método de ajuste Propuesta.

Los polinomios son funciones comúnmente usadas para aproximar datos u otras funciones que no se pueden recrear en la práctica. Un ejemplo conocido es el hecho de que las calculadoras científicas no trabajan con la función seno(x) si no con polinomios que la aproximan.

$$\text{seno}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

El método consta de dos pasos principales:

1. Encontrar una función polinomio que se ajuste a los datos  $-\ln(p_x)$  de la tabla de mortalidad.
2. Transformar el polinomio del paso anterior en uno que ajuste a la fuerza de mortalidad.

### 4.1 Polinomios.

#### Definición.

Se dice que una función  $P(x)$  es un polinomio si se puede escribir de la siguiente manera:

$$P(x) = a_0 + a_1x^1 + a_2x^2 + \dots + a_nx^n, \text{ donde } a_i \text{ es constante } \forall i \in \mathbb{N}, 0 \leq i \leq n$$

Apoyándonos en ésta definición podemos decir que la función  $f(x) = 1 + 2x + x^2$  es un polinomio, así como también lo es  $g(x) = (1 - x)(1 + x)$  pues ésta última puede ser escrita como  $g(x) = 1 - x^2$ . Por otra parte, la función  $h(x) = 1 + x^{-1}$  no es de tipo polinomio por que  $x$  esta elevada a la potencia  $-1$ .

#### Grado del Polinomio.

Se define como Grado del polinomio, generalmente denotado por una  $n$ , al número natural dado por:

$$n = \{\text{máx. } (k) \text{ } \forall a_k \neq 0\}.$$

Por ejemplo, el polinomio  $P(x) = 3x^4 + 5x + 1$  es de grado 4, mientras que  $P(x) = 5$  es de grado cero.

Se define como Polinomio Cero o Polinomio Nulo, a aquel en el que todas las constantes del polinomio  $a_i$  son cero, es decir,  $P(x) = 0$ . Nótese que el polinomio cero no tiene grado definido, lo cual es totalmente diferente a un polinomio de grado cero.

### **Dominio y Rango de los Polinomios.**

Como se desprende de la definición de polinomio, en  $P(x)$  no existen restricciones para los valores que puede tomar la variable independiente  $x$ . Por lo tanto el Dominio para cualquier función polinomio es el conjunto de los números reales.

Así mismo, si el grado del polinomio es impar, entonces el Rango del polinomio será el conjunto de los números reales. Por otra parte, si el polinomio es de grado par, entonces el Rango del polinomio estará conformado por un subconjunto de los números reales. Por ejemplo,  $P(x) = 1 - x^2$  tiene como Rango el conjunto de los números reales menores o iguales a uno.

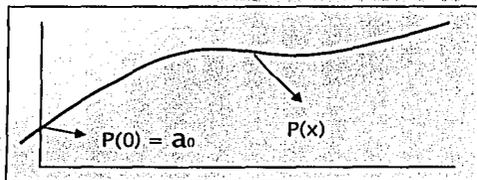
Usando la notación de función como regla de correspondencia entre dos conjuntos, todo lo anteriormente dicho en este inciso se puede resumir mediante la siguiente expresión:

$P(x): \mathbb{R} \rightarrow B \subset \mathbb{R}$  si  $P(x)$  es de grado par y

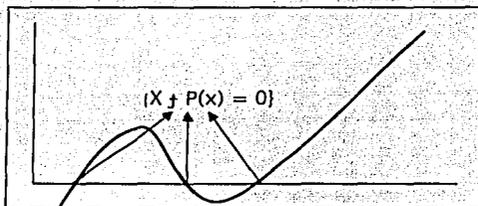
$P(x): \mathbb{R} \rightarrow \mathbb{R}$  si  $P(x)$  es de grado impar.

## Intersecciones con los ejes.

El valor  $P(0) = a_0$  se conoce como ordenada al origen y es el punto donde la gráfica de  $P(x)$  intercepta al eje de las ordenadas:



Las soluciones de la ecuación  $P(x) = 0$  se conocen como Raíces del Polinomio y es el conjunto de Intersecciones de  $P(x)$  con el eje de las abscisas. El conjunto  $\{x_i \mid P(x_i) = 0\}$  puede ser vacío para polinomios de grado par:

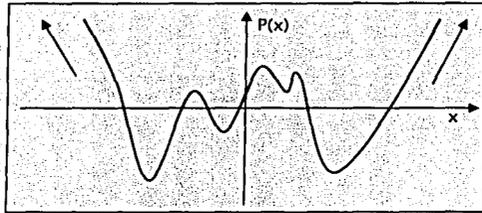


Sin importar de qué polinomio se trate, siempre y cuando sea de grado impar, se puede asegurar que al menos existe una raíz.

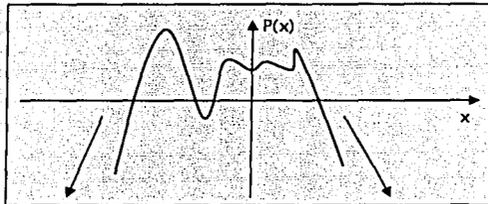
## Término Dominante.

Sea  $n$  el grado del polinomio  $P(X) = a_0 + a_1X + a_2X^2 + \dots + a_nX^n$ , el *término dominante* está dado por la constante  $a_n$ . La importancia de dicha constante es que determinan el comportamiento de la función cuando  $x \rightarrow \pm\infty$ .

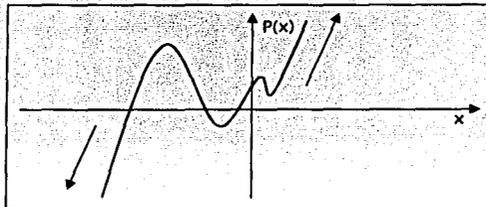
Si el grado del polinomio  $n$  es par y el término dominante  $a_0$  es positivo,  $P(x) \rightarrow \infty$  cuando  $x \rightarrow \pm\infty$



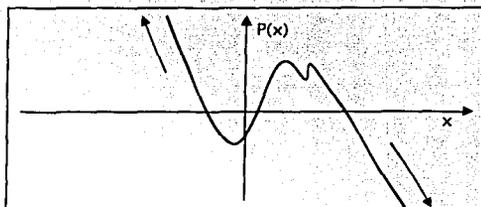
Si el grado del polinomio  $n$  es par y el término dominante  $a_0$  es negativo,  $P(x) \rightarrow -\infty$  cuando  $x \rightarrow \pm\infty$



Si el grado del polinomio  $n$  es impar y el término dominante  $a_0$  es positivo,  $P(x) \rightarrow \infty$  cuando  $x \rightarrow \infty$  y  $P(x) \rightarrow -\infty$  cuando  $x \rightarrow -\infty$ .



Si el grado del polinomio  $n$  es impar y el término dominante  $a_0$  es negativo,  $P(x) \rightarrow -\infty$  cuando  $x \rightarrow \infty$  y  $P(x) \rightarrow \infty$  cuando  $x \rightarrow -\infty$ .



### Derivación e Integración.

Todo polinomio, sin importar su grado, tiene función derivada y función primitiva en el intervalo de los números reales. Además, la derivada y la primitiva de un polinomio es a su vez un polinomio.

### Polinomios y Espacios Vectoriales.

El conjunto de funciones polinomios es un espacio Vectorial. Los elementos de los arreglos (o vectores) están conformados por las constantes  $a_0, a_1, \dots, a_n$ , y no así por los valores que puede tomar la función.

Puesto que dicho espacio vectorial es de dimensión infinita numerable, lo cual no tiene ninguna utilidad práctica, se suele utilizar un subespacio vectorial de dimensión finita. Este subespacio se define como el conjunto de polinomios de grado menor o igual a  $n$ .

La utilidad del hecho de que los polinomios sean espacios vectoriales, es que podemos representar transformaciones lineales entre estos mediante el empleo de matrices, y la composición de transformaciones efectuarlas mediante el producto de matrices lo cual utilizaremos en el presente capítulo.

## 4.2 Método de Ajuste.

### Supuesto.

La hipótesis principal es que la fuerza de mortalidad puede ser adecuadamente representada mediante un polinomio, es decir:

$$\mu_x \approx P(x) = a_0 + a_1x^1 + a_2x^2 + \dots + a_nx^n \quad \dots (73)$$

### Matriz de Relación entre los polinomios de $\mu_x$ y $p_x$ .

El primer paso es establecer una igualdad que relacione las tasas de mortalidad con la fuerza de mortalidad<sup>(4)</sup>. Una igualdad del Capítulo 2 puede ser nuestro punto de partida:

$${}_t p_x = e^{-\int_x^{x+t} \mu_s ds} \quad \forall t \geq 0$$

$$\Rightarrow p_x = e^{-\int_x^{x+1} \mu_s ds}$$

Con la intención de despejar la integral, aplicamos logaritmo natural y multiplicando por (-1):

$$-\ln(p_x) = \int_x^{x+1} \mu_s ds$$

Ahora es cuando aplicamos el supuesto (73):

$$\Rightarrow -\ln(p_x) = \int_x^{x+1} (a_0 + a_1s + a_2s^2 + \dots + a_ns^n) ds$$

$$\Rightarrow -\ln(p_x) = \left[ a_0 \cdot s + a_1 \frac{s^2}{2} + \dots + a_n \frac{s^{n+1}}{n+1} \right]_x^{x+1}$$

$$\Rightarrow -\ln(p_x) = \left[ a_0(x+1) + a_1 \frac{(x+1)^2}{2} + \dots + a_n \frac{(x+1)^{n+1}}{n+1} \right] - \left[ a_0x + a_1 \frac{(x)^2}{2} + \dots + a_n \frac{(x)^{n+1}}{n+1} \right]$$

<sup>(4)</sup> Se utiliza  $t=1$  debido a que las tablas de mortalidad son por periodos anuales. El parámetro "t" se puede generalizar dependiendo del intervalo de medición de las tablas.

Con base en la fórmula del Binomio de Newton  $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$ , podemos desarrollar las potencias  $(x+1)^k$  como sumas que nos permitirán simplificar más adelante:

$$\Rightarrow -\ln(p_x) = \left[ a_0(x+1) + a_1 \frac{\sum_{k=0}^2 \binom{2}{k} x^{2-k}}{2} + \dots + a_n \frac{\sum_{k=0}^{n+1} \binom{n+1}{k} x^{n+1-k}}{n+1} \right] - \left[ a_0 x + a_1 \frac{(x)^2}{2} + \dots + a_n \frac{(x)^{n+1}}{n+1} \right]$$

El primer término de cada sumatoria se puede simplificar como  $\frac{\binom{i}{0} x^{i-0}}{i} = \frac{x^i}{i}$ , lo cual nos permite sustraer su término semejante

$$\Rightarrow -\ln(p_x) = \left[ a_0 \cancel{(x+1)} + a_1 \frac{\sum_{k=0}^2 \binom{2}{k} x^{2-k}}{2} + \dots + a_n \frac{\sum_{k=0}^{n+1} \binom{n+1}{k} x^{n+1-k}}{n+1} \right] - \left[ a_0 \cancel{x} + a_1 \cancel{\frac{(x)^2}{2}} + \dots + a_n \cancel{\frac{(x)^{n+1}}{n+1}} \right]$$

y reducir la expresión anterior en:

$$\Rightarrow -\ln(p_x) = \left[ a_0 + a_1 \frac{\sum_{k=1}^2 \binom{2}{k} x^{2-k}}{2} + \dots + a_n \frac{\sum_{k=1}^{n+1} \binom{n+1}{k} x^{n+1-k}}{n+1} \right]$$

Desarrollando términos semejantes y factorizando en términos de  $x!$ :

$$\begin{aligned} \Rightarrow -\ln(p_x) &= \left( a_0 + \frac{a_1}{2} \binom{2}{2} + \frac{a_2}{3} \binom{3}{3} + \dots + \frac{a_{n-1}}{n} \binom{n}{n} + \frac{a_n}{n+1} \binom{n+1}{n+1} \right) x^0 + \\ &+ \left( \frac{a_1}{2} \binom{2}{2-1} + \frac{a_2}{3} \binom{3}{3-1} + \dots + \frac{a_{n-1}}{n} \binom{n}{n-1} + \frac{a_n}{n+1} \binom{n+1}{(n+1)-1} \right) x^1 + \\ &\vdots \\ &+ \left( \frac{a_{n-1}}{n} \binom{n}{n-(n-1)} + \frac{a_n}{n+1} \binom{n+1}{(n+1)-n-1} \right) x^{n-1} + \\ &+ \left( \frac{a_n}{n+1} \binom{n+1}{(n+1)-n} \right) x^n \end{aligned}$$

Se puede demostrar fácilmente del cálculo combinatorio que  $\binom{n}{n-k} = \binom{n}{k}$  (ver anexo). Así se simplifica nuestra expresión en:

$$\begin{aligned} \Rightarrow -\ln(p_x) &= \left( a_0 + \frac{a_1}{2} \binom{2}{0} + \frac{a_2}{3} \binom{3}{0} + \dots + \frac{a_{n-1}}{n} \binom{n}{0} + \frac{a_n}{n+1} \binom{n+1}{0} \right) x^0 + \\ &+ \left( \frac{a_1}{2} \binom{2}{1} + \frac{a_2}{3} \binom{3}{1} + \dots + \frac{a_{n-1}}{n} \binom{n}{1} + \frac{a_n}{n+1} \binom{n+1}{1} \right) x^1 + \\ &\vdots \\ &+ \left( \frac{a_{n-1}}{n} \binom{n}{n-1} + \frac{a_n}{n+1} \binom{n+1}{n-1} \right) x^{n-1} + \\ &+ \left( \frac{a_n}{n+1} \binom{n+1}{n} \right) x^n \end{aligned}$$

Si se observa con cuidado el extremo derecho de la igualdad anterior, se llegó a un nuevo polinomio en el que cada término  $a_i$  está dado por la suma de términos entre corchetes que multiplica a cada  $x^i$ :

$$\therefore -\ln(p_x) = \bar{a}_0 + \bar{a}_1 x + \dots + \bar{a}_n x^n \quad (74)$$

Este resultado es muy importante, pues con él se obtienen los dos principales soportes para nuestro método:

1. Si  $\mu_x$  se puede determinar o aproximar mediante un polinomio ( $\mu_x \approx P(x) = a_0 + a_1 x + \dots + a_n x^n$ ), entonces  $-\ln(p_x)$  también puede ser expresado mediante otro polinomio ( $-\ln(p_x) = \bar{a}_0 + \bar{a}_1 x + \dots + \bar{a}_n x^n$ ).

2. Los términos  $(a_0, a_1, \dots, a_n)$  del polinomio para  $\mu$ , y los términos  $(\bar{a}_0, \bar{a}_1, \dots, \bar{a}_n)$  del polinomio para  $-\ln(p_x)$  están relacionados entre sí, y es posible encontrar a los coeficientes en términos de los otros:

$$\left\{ \begin{array}{l} \bar{a}_0 = a_0 + \frac{a_1}{2} \binom{2}{0} + \frac{a_2}{3} \binom{3}{0} + \dots + \frac{a_{n-1}}{n} \binom{n}{0} + \frac{a_n}{n+1} \binom{n+1}{0} \\ \bar{a}_1 = \frac{a_1}{2} \binom{2}{1} + \frac{a_2}{3} \binom{3}{1} + \dots + \frac{a_{n-1}}{n} \binom{n}{1} + \frac{a_n}{n+1} \binom{n+1}{1} \\ \vdots \\ \bar{a}_{n-1} = \frac{a_{n-1}}{n} \binom{n}{n-1} + \frac{a_n}{n+1} \binom{n+1}{n-1} \\ \bar{a}_n = \frac{a_n}{n+1} \binom{n+1}{n} \end{array} \right. \quad (75)$$

El sistema de ecuaciones obtenido también puede ser escrito mediante matrices cuadradas (transformaciones) y vectores:

$$\begin{pmatrix} \bar{a}_0 \\ \bar{a}_1 \\ \vdots \\ \bar{a}_{n-1} \\ \bar{a}_n \end{pmatrix} = \begin{pmatrix} 1 \binom{1}{0} & \frac{1}{2} \binom{2}{0} & \frac{1}{3} \binom{3}{0} & \dots & \frac{1}{n} \binom{n}{0} & \frac{1}{n+1} \binom{n+1}{0} \\ & \frac{1}{2} \binom{2}{1} & \frac{1}{3} \binom{3}{1} & \dots & \frac{1}{n} \binom{n}{1} & \frac{1}{n+1} \binom{n+1}{1} \\ & & \ddots & & & \vdots \\ & & & & \frac{1}{n} \binom{n}{n-1} & \frac{1}{n+1} \binom{n+1}{n-1} \\ & & & & & \frac{1}{n+1} \binom{n+1}{n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} \quad (76)$$

es decir,

$$\bar{\mathbf{a}} = \mathbf{A} \cdot \mathbf{a}$$

donde

$$\bar{\mathbf{a}} = (\bar{a}_0, \bar{a}_1, \dots, \bar{a}_n)$$

$$\mathbf{a} = (a_0, a_1, \dots, a_n)$$

$$A = \begin{bmatrix} 1 \binom{1}{0} & \frac{1}{2} \binom{2}{0} & \frac{1}{3} \binom{3}{0} & \dots & \frac{1}{n} \binom{n}{0} & \frac{1}{n+1} \binom{n+1}{0} \\ & \frac{1}{2} \binom{2}{1} & \frac{1}{3} \binom{3}{1} & \dots & \frac{1}{n} \binom{n}{1} & \frac{1}{n+1} \binom{n+1}{1} \\ & & \ddots & & & \vdots \\ & & & & \frac{1}{n} \binom{n}{n-1} & \frac{1}{n+1} \binom{n+1}{n-1} \\ & & & & & \frac{1}{n+1} \binom{n+1}{n} \end{bmatrix} \quad (77)$$

Podemos concluir que, la matriz A sirve para encontrar al polinomio de  $\mu_x$  en términos del polinomio de  $-\ln(p_x)$  y viceversa.

$$\Rightarrow A^{-1} \begin{pmatrix} \bar{a}_0 \\ \bar{a}_1 \\ \vdots \\ \bar{a}_{n-1} \\ \bar{a}_n \end{pmatrix} = (A^{-1} \cdot A) \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} \quad (78)$$

Por ser A una matriz de tipo triangular superior en la que todos los elementos de la diagonal son distintos de cero, podemos asegurar que  $A^{-1}$  existe. A la matriz A se le puede escribir de manera simplificada mediante la siguiente regla o algoritmo:

$$a_{i,j} = \begin{cases} \frac{1}{j} \binom{j}{i-1} & \text{si } i \leq j \\ 0 & \text{si } i > j \end{cases}, \quad a_{i,j} \in m_{n+1 \times n+1} \quad (79)$$

### Cálculo del polinomio que ajusta a $-\ln(p_x)$ .

El siguiente paso en nuestro método de ajuste es determinar los coeficientes  $a'_0, a'_1, \dots, a'_n$  del polinomio que ajusta a  $-\ln(p_x)$ .

Existen diversos procedimientos ya establecidos, que por lo general se muestran en libros de análisis numérico, que permiten estimar una serie de datos mediante una función polinomio. No obstante cualquiera de ellos puede aplicarse en nuestro problema, se recomienda el denominado mínimos cuadrados por matrices transpuestas.

El procedimiento de mínimos cuadrados además de ser estandarizable, tiene la ventaja de resultar en un polinomio lo más sencillo posible minimizando el error entre el modelo y los datos. Es deseable que el grado del polinomio sea lo más pequeño posible, pues el tamaño de la matriz  $A$  depende de este y entre más grande sea  $A$  más difícil será calcular  $A^{-1}$ .

Otros procedimientos como por ejemplo Base de Newton, resultan en un polinomio que pasa exactamente por los puntos a estimar, pero tienen la desventaja que el polinomio obtenido es demasiado complejo.

### Uso de mínimos cuadrados para determinar $\tilde{a}_0, \tilde{a}_1, \dots, \tilde{a}_n$ .

Tomemos como base la siguiente tabla o fracción de tabla de mortalidad:

$x$ (Edad)	$q_x$
$x_1$	$q_{x_1}$
$x_2$	$q_{x_2}$
...	...
$x_k$	$q_{x_k}$

donde:  $1 \leq x_1 < x_k \leq w-1$ .

Como se explicó brevemente, el procedimiento de mínimos cuadrados funciona y presupone la existencia de un error entre los datos y la función que los aproxima. La diferencia entre los datos ( $-\ln(p_x)$ ) y el polinomio de aproximación ( $\tilde{a}_0 + \tilde{a}_1x + \dots + \tilde{a}_nx^n$ ) para cada una de las edades genera una tabla de errores. Esa tabla es la base para introducir y solucionar el problema por el método de mínimos cuadrados.

x (Edad)	$y = -\ln(1 - q_x)$	$\hat{y} = \bar{a}_0 + \bar{a}_1 x + \dots + \bar{a}_n x^n$	$\epsilon_i^2 = [y - \hat{y}]^2$
$x_1$	$-\ln(1 - q_{x_1})$	$\bar{a}_0 + \bar{a}_1 x_1 + \dots + \bar{a}_n x_1^n$	$[-\ln(1 - q_{x_1}) - (\bar{a}_0 + \bar{a}_1 x_1 + \dots + \bar{a}_n x_1^n)]^2$
$x_2$	$-\ln(1 - q_{x_2})$	$\bar{a}_0 + \bar{a}_1 x_2 + \dots + \bar{a}_n x_2^n$	$[-\ln(1 - q_{x_2}) - (\bar{a}_0 + \bar{a}_1 x_2 + \dots + \bar{a}_n x_2^n)]^2$
...	...	...	...
$x_k$	$-\ln(1 - q_{x_k})$	$\bar{a}_0 + \bar{a}_1 x_k + \dots + \bar{a}_n x_k^n$	$[-\ln(1 - q_{x_k}) - (\bar{a}_0 + \bar{a}_1 x_k + \dots + \bar{a}_n x_k^n)]^2$

A partir del objetivo de minimizar la suma de los errores  $\epsilon_i^2$  de la tabla anterior, es que podemos aplicar el procedimiento de los mínimos cuadrados. No se pretende el desarrollo o justificación de la fórmula de mínimos cuadrados para minimizar los residuos, pues esta puede ser encontrada en cualquier libro de análisis numérico o estadística:

$$\alpha = (X^t \cdot X)^{-1} \cdot X^t \cdot Y \quad \dots (80)$$

donde

$$\alpha = \begin{pmatrix} \bar{a}_0 \\ \bar{a}_1 \\ \vdots \\ \bar{a}_n \end{pmatrix} \in m_{n \times 1}$$

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & x_k^2 & \dots & x_k^n \end{pmatrix} \in m_{k \times n}$$

$$Y = \begin{pmatrix} -\ln(1 - q_{x_1}) \\ -\ln(1 - q_{x_2}) \\ \vdots \\ -\ln(1 - q_{x_k}) \end{pmatrix} \in m_{k \times 1}$$

Obsérvese que precisamente los elementos del vector  $\alpha$  son los coeficientes del polinomio  $-\ln(p_x) \approx \bar{a}_0 + \bar{a}_1x + \dots + \bar{a}_nx^n$ . Por lo tanto el siguiente y último paso es calcular el polinomio  $\mu_x = a_0 + a_1x + \dots + a_nx^n$ .

### Cálculo de $\mu_x$ .

Para concluir con el presente método solo nos resta encontrar los términos del polinomio  $\mu_x = a_0 + a_1x + \dots + a_nx^n$ .

Si recordamos, en secciones anteriores vimos que esto es posible gracias a la matriz A que define la siguiente relación:

$$\begin{pmatrix} \bar{a}_0 \\ \bar{a}_1 \\ \vdots \\ \bar{a}_n \end{pmatrix} = A \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = A^{-1} \cdot \begin{pmatrix} \bar{a}_0 \\ \bar{a}_1 \\ \vdots \\ \bar{a}_n \end{pmatrix}$$

Donde:

$$\mu_x = a_0 + a_1x + \dots + a_nx^n$$

$$-\ln(p_x) \approx \bar{a}_0 + \bar{a}_1x + \dots + \bar{a}_nx^n$$

y los elementos de la matriz A definidos por:

$$a_{i,j} = \begin{cases} \frac{1}{j} \binom{j}{i} & \text{si } i \leq j \\ 0 & \text{si } i > j \end{cases}, a_{i,j} \in m_{n+1 \times n+1}$$

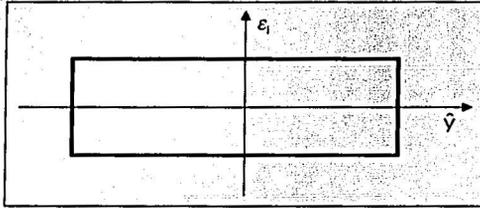
La matriz inversa  $A^{-1}$  se puede obtener mediante paquetería como Matlab o Excel. Otra forma de determinar los términos del  $\mu_x = a_0 + a_1x + \dots + a_nx^n$  es resolviendo directamente el sistema de ecuaciones (75), despejando el valor de  $a_n$  en la última ecuación del sistema y substituirlo en la ecuación inmediata superior para a su vez poder despejar  $a_{n-1}$ . Repitiendo este procedimiento es posible encontrar todos los términos.

## Control del Error Relativo con Mínimos Cuadrados con Factores de Peso.

Existen diversos métodos de análisis para medir la confiabilidad de un modelo que estima una serie de datos. El más importante y el único que en este caso nos resulta concluyente, es el análisis de los residuos.

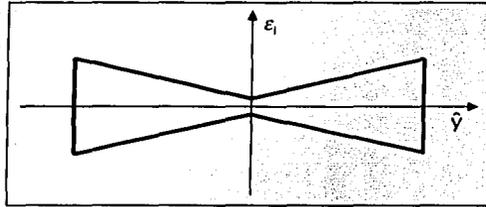
El método de mínimos cuadrados tradicional funciona bajo el supuesto que la esperanza del error es cero y que la varianza es constante, es decir, se busca que el error sea homocedástico  $\varepsilon_i \sim N(0, \sigma^2)$ .

El rectángulo que se muestra en la siguiente figura, ejemplifica este supuesto de los mínimos cuadrados tradicional.



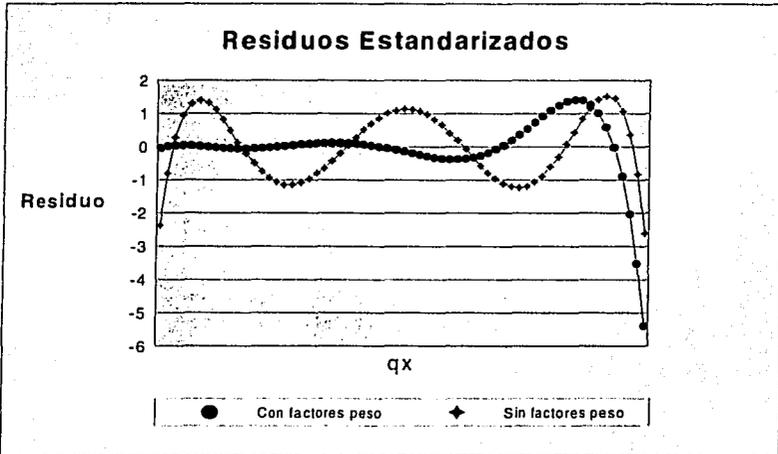
Sin embargo, nosotros requerimos de un modelo que proporcione resultados más precisos cuando las tasas de mortalidad son pequeñas y resultados más holgados cuando las tasas de mortalidad son más grandes. Es decir, nuestro caso en particular requiere que la varianza del error no sea constante, si no más bien, sea de alguna manera proporcional al valor de  $q_x$  para cada edad. El trasfondo de este requerimiento viene de un principio de análisis numérico en el que el error relativo entre los datos y el modelo debe mantenerse constante.

Un error relativo constante ocasiona que los residuos del modelo sean proporcionales a la variable de predicción como se muestra a continuación:



Precisamente en casos en que la varianza de los residuos no se mantiene constante es que se debe aplicar *mínimos cuadrados con factores de peso*

En la siguiente gráfica se muestra como el aplicar los factores de peso ocasiona el cambio deseado en los residuos y en el error relativo:



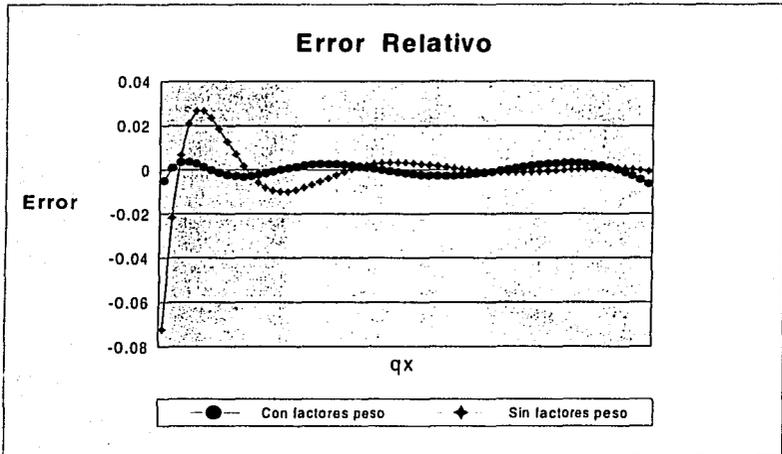
Donde:

$$\text{Residuo Estandarizado} = \epsilon_i / s$$

$$\epsilon_i = i\text{-ésimo residuo} = y_i - \hat{y}_i$$

$$s = \text{desviación estándar residual} = \sqrt{\text{MCE}}$$

$$\text{MCE} = \text{Desviación estándar residual} = \frac{\sum_{i=1}^n (\epsilon_i^2)}{n-2}$$



Donde :

$$\text{Error relativo} = \frac{(y_i - \hat{y}_i)}{y_i}$$

La manera de aplicar mínimos cuadrados con factores de peso es casi idéntica a la manera tradicional, solo que se integra al modelo la matriz de factores de peso Q:

$$\alpha = (X^t \cdot Q^{-1} \cdot X) \cdot X^t \cdot Q^{-1} \cdot Y \quad \dots (81)$$

Donde:

$\alpha$ , X y Y son los definidos en (80),

Q es la matriz de factores de peso llamada matriz de ponderación. Q es de tipo diagonal, la cual tradicionalmente contiene la varianza de cada una de los residuos en la diagonal y ceros en las posiciones no diagonales de la matriz, es decir:

$$Q = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}$$

$$\Rightarrow Q^{-1} = \begin{pmatrix} 1/\sigma_1^2 & 0 & \dots & 0 \\ 0 & 1/\sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/\sigma_n^2 \end{pmatrix}$$

Nótese que si  $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$  entonces  $Q = \sigma^2 \cdot I_n$  con lo que (81) se reduce a (80).

### Definición de la matriz de factores de peso.

Se proponen dos maneras diferentes de llenar los elementos  $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$  de la matriz Q.

Podemos definir una variable aleatoria muy sencilla para cada edad x que nos ayude a determinar una varianza diferente para cada edad. Sea la variable aleatoria  $I_{(x)}$  definida por

$$I_{(x)} = \begin{cases} 1 & \text{Si } (x) \text{ muere antes de la edad } x+1 \\ 0 & \text{e.o.c.} \end{cases}$$

La distribución de  $I_{(x)}$  es de tipo Bernoulli, por lo que la media estará dada por  $\mu_{I_{(x)}} = E[I_{(x)}] = q_x$  y la varianza por  $\sigma^2 = q_x p_x$ . Entonces la matriz de factores de pesos quedaría de la siguiente manera:

$$Q = \begin{pmatrix} q_{x_1} p_{x_1} & 0 & \dots & 0 \\ 0 & q_{x_2} p_{x_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & q_{x_n} p_{x_n} \end{pmatrix} \quad \dots(82)$$

Otra manera de proponer la matriz Q es desde el punto de vista de análisis numérico. Si proponemos que la dispersión del cada error  $\epsilon_i$  sea proporcional a la

magnitud de las tasas de mortalidad para cada edad, obtenemos que la desviación estándar para cada error  $\varepsilon_i$  es proporcional a  $q_x$ :

$$\sigma_i = k \cdot q_x$$

Por lo que la matriz Q estará dada por:

$$Q = \begin{pmatrix} (k \cdot q_{x_1})^2 & 0 & \dots & 0 \\ 0 & (k \cdot q_{x_2})^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (k \cdot q_{x_n})^2 \end{pmatrix}$$

$$\Rightarrow Q^{-1} = \begin{pmatrix} 1/(k \cdot q_{x_1})^2 & 0 & \dots & 0 \\ 0 & 1/(k \cdot q_{x_2})^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/(k \cdot q_{x_n})^2 \end{pmatrix}$$

El aplicar  $Q^{-1}$  o  $k^2 \cdot Q^{-1}$  en (81) es totalmente equivalente, por lo que podremos utilizar:

$$Q^{-1} = \begin{pmatrix} 1/(q_{x_1})^2 & 0 & \dots & 0 \\ 0 & 1/(q_{x_2})^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/(q_{x_n})^2 \end{pmatrix} \quad \dots (83)$$

### 4.3 Aplicación.

Para ejemplo se utilizó la Tabla de Mortalidad "CNSF 2000 - I (1991-1998)" en el rango de edades 13 a 99 años. Utilizaremos un polinomio de grado 5, es decir, supondremos que existen  $a_0, a_1, \dots, a_5$  tales que  $\mu_x = a_0 + a_1x + \dots + a_5x^5$ . Aplicando la metodología previamente desarrollada para las edades [13,99] tenemos:

$$X = \begin{pmatrix} 1 & x_1 & \dots & x_1^5 \\ 1 & x_2 & \dots & x_2^5 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{88} & \dots & x_{88}^5 \end{pmatrix} = \begin{pmatrix} 1 & 13 & \dots & 13^5 \\ 1 & 14 & \dots & 14^5 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 99 & \dots & 99^5 \end{pmatrix}$$

$$\Rightarrow X^t = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 13 & 14 & \dots & 99 \\ \vdots & \vdots & \ddots & \vdots \\ 13^5 & 14^5 & \dots & 99^5 \end{pmatrix}$$

$$Y = \begin{pmatrix} -\ln(1-q_{x_1}) \\ -\ln(1-q_{x_2}) \\ \vdots \\ -\ln(1-q_{x_n}) \end{pmatrix} = \begin{pmatrix} -\ln(1-q_{13}) \\ -\ln(1-q_{14}) \\ \vdots \\ -\ln(1-q_{99}) \end{pmatrix} = \begin{pmatrix} 0.000427091 \\ 0.000460106 \\ \vdots \\ 0.22795134 \end{pmatrix}$$

$$Q = \begin{pmatrix} (q_{x_1})^2 & 0 & \dots & 0 \\ 0 & (q_{x_2})^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (q_{x_n})^2 \end{pmatrix} = \begin{pmatrix} (q_{13})^2 & 0 & \dots & 0 \\ 0 & (q_{14})^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (q_{99})^2 \end{pmatrix}$$

$$= \begin{pmatrix} 1.823290 \times 10^{-07} & 0 & \dots & 0 \\ 0 & 2.116000 \times 10^{-07} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 4.154952 \times 10^{-02} \end{pmatrix}$$

$$\Rightarrow Q^{-1} = \begin{pmatrix} \frac{1}{1.823290 \times 10^{-07}} & 0 & \dots & 0 \\ 0 & \frac{1}{2.116000 \times 10^{-07}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{4.154952 \times 10^{-02}} \end{pmatrix}$$

Si se desea, se pueden verificar los valores de  $q_x$  en la tabla del capítulo 2. Ahora si hacemos uso de (81) tenemos que:

$$\alpha = (X^t Q^{-1} X)^{-1} \cdot X^t \cdot Q^{-1} \cdot Y$$

Haciendo uso del mismo procedimiento que para mínimos cuadrados ordinarios se llega a que:

$$\Rightarrow \alpha = \begin{pmatrix} \bar{a}_0 \\ \bar{a}_1 \\ \bar{a}_2 \\ \bar{a}_3 \\ \bar{a}_4 \\ \bar{a}_5 \end{pmatrix} = \begin{pmatrix} -1.611240 \times 10^{-03} \\ 3.582164 \times 10^{-04} \\ -2.48969 \times 10^{-05} \\ 8.85985 \times 10^{-07} \\ -1.42562 \times 10^{-08} \\ 9.88911 \times 10^{-11} \end{pmatrix}$$

Si recordamos la finalidad del vector  $\alpha$ , tan solo hemos encontrado las constantes  $\bar{a}_0, \bar{a}_1, \dots, \bar{a}_5$  que determinan al polinomio para  $-\ln(p_x)$ . Ahora con ayuda de la matriz A transformaremos a  $\alpha$  en el polinomio que aproxima a  $\mu_x$ . Si el polinomio es de grado 5 y A está detallado por (77) tenemos que:

$$A_5 = \begin{pmatrix} 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \frac{1}{2} \begin{pmatrix} 2 \\ 0 \end{pmatrix} & \frac{1}{3} \begin{pmatrix} 3 \\ 0 \end{pmatrix} & \frac{1}{4} \begin{pmatrix} 4 \\ 0 \end{pmatrix} & \frac{1}{5} \begin{pmatrix} 5 \\ 0 \end{pmatrix} & \frac{1}{6} \begin{pmatrix} 6 \\ 0 \end{pmatrix} \\ 0 & \frac{1}{2} \begin{pmatrix} 2 \\ 1 \end{pmatrix} & \frac{1}{3} \begin{pmatrix} 3 \\ 1 \end{pmatrix} & \frac{1}{4} \begin{pmatrix} 4 \\ 1 \end{pmatrix} & \frac{1}{5} \begin{pmatrix} 5 \\ 1 \end{pmatrix} & \frac{1}{6} \begin{pmatrix} 6 \\ 1 \end{pmatrix} \\ 0 & 0 & \frac{1}{3} \begin{pmatrix} 3 \\ 2 \end{pmatrix} & \frac{1}{4} \begin{pmatrix} 4 \\ 2 \end{pmatrix} & \frac{1}{5} \begin{pmatrix} 5 \\ 2 \end{pmatrix} & \frac{1}{6} \begin{pmatrix} 6 \\ 2 \end{pmatrix} \\ 0 & 0 & 0 & \frac{1}{4} \begin{pmatrix} 4 \\ 3 \end{pmatrix} & \frac{1}{5} \begin{pmatrix} 5 \\ 3 \end{pmatrix} & \frac{1}{6} \begin{pmatrix} 6 \\ 3 \end{pmatrix} \\ 0 & 0 & 0 & 0 & \frac{1}{5} \begin{pmatrix} 5 \\ 4 \end{pmatrix} & \frac{1}{6} \begin{pmatrix} 6 \\ 4 \end{pmatrix} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} \begin{pmatrix} 6 \\ 5 \end{pmatrix} \end{pmatrix}$$

$$\Rightarrow A_5 = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ 0 & \frac{2}{2} & \frac{3}{3} & \frac{4}{4} & \frac{5}{5} & \frac{6}{6} \\ 0 & 0 & \frac{3}{3} & \frac{6}{4} & \frac{10}{5} & \frac{15}{6} \\ 0 & 0 & 0 & \frac{4}{4} & \frac{10}{5} & \frac{20}{6} \\ 0 & 0 & 0 & 0 & \frac{5}{5} & \frac{15}{6} \\ 0 & 0 & 0 & 0 & 0 & \frac{6}{6} \end{pmatrix}$$

$$\Rightarrow (A_5)^{-1} = \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{6} & 0 & -\frac{1}{30} & 0 \\ 0 & 1 & -1 & \frac{1}{2} & 0 & -\frac{1}{6} \\ 0 & 0 & 1 & -\frac{3}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & \frac{5}{3} \\ 0 & 0 & 0 & 0 & 1 & -\frac{5}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Si aplicamos  $A^{-1}$  como transformación al vector  $\alpha$  encontramos que:

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = (A_5)^{-1} \begin{pmatrix} \tilde{a}_0 \\ \tilde{a}_1 \\ \tilde{a}_2 \\ \tilde{a}_3 \\ \tilde{a}_4 \\ \tilde{a}_5 \end{pmatrix} = \begin{pmatrix} -1.794498 \times 10^{-03} \\ 3.835563 \times 10^{-04} \\ -2.624015 \times 10^{-05} \\ 9.146621 \times 10^{-07} \\ -1.450343 \times 10^{-08} \\ 9.889105 \times 10^{-11} \end{pmatrix}$$

Es decir:

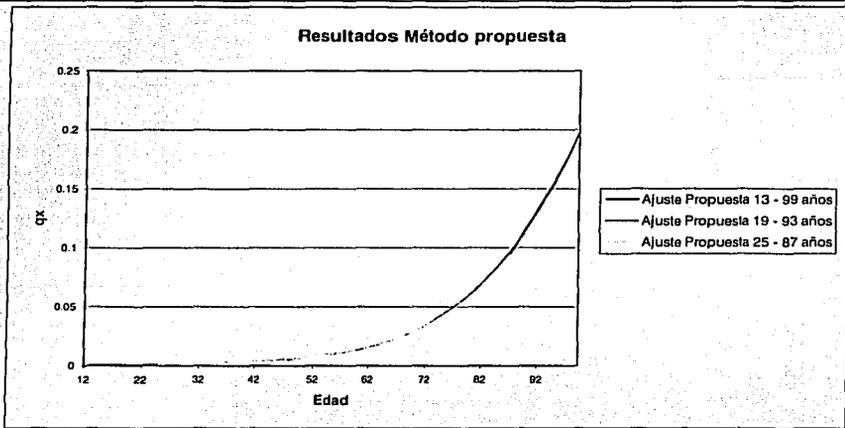
$$\mu_x = -1.794498 \times 10^{-03} + (3.835563 \times 10^{-04})x + (-2.624015 \times 10^{-05})x^2 + (9.146621 \times 10^{-07})x^3 + (-1.450343 \times 10^{-08})x^4 + (9.889105 \times 10^{-11})x^5$$

## 4.4 Resultados.

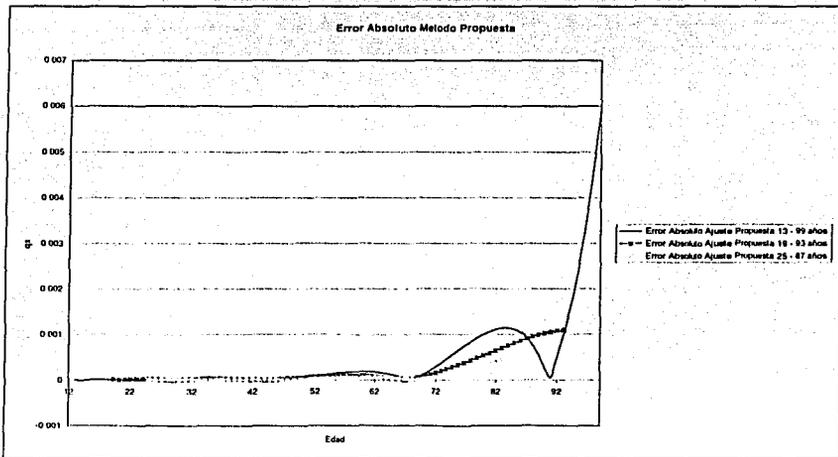
Para observar la bondad de ajuste entre Makeham (primera Ley) contra el polinomio (metodología propuesta), se aplicó el método tres veces para los diferentes rangos de edades [13,99], [19,93] y [25,87] años. Los resultados se muestran en la siguiente tabla:

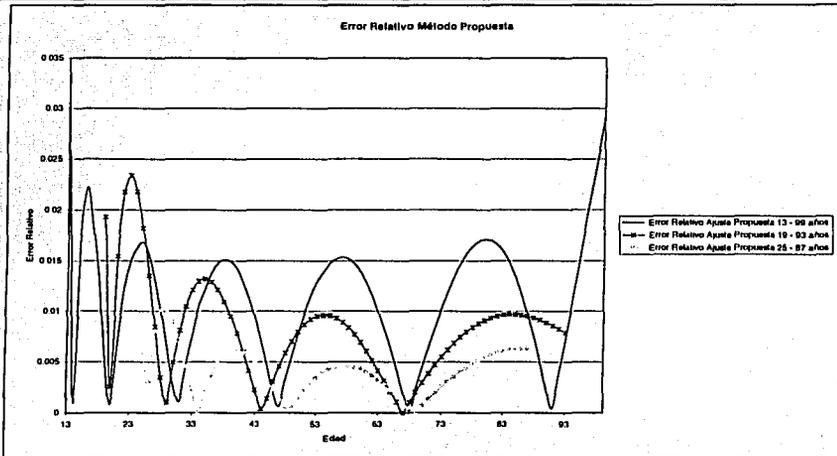
	[13,99] años.	[19,93] años.	[25,87] años.
$a_0$	$-1.794498 \times 10^{-03}$	$-4.603132 \times 10^{-03}$	$-9.136361 \times 10^{-03}$
$a_1$	$3.835563 \times 10^{-04}$	$7.479708 \times 10^{-04}$	$1.253224 \times 10^{-03}$
$a_2$	$-2.624015 \times 10^{-05}$	$-4.371653 \times 10^{-05}$	$-6.502154 \times 10^{-05}$
$a_3$	$9.146621 \times 10^{-07}$	$1.300139 \times 10^{-06}$	$1.723230 \times 10^{-06}$
$a_4$	$-1.450343 \times 10^{-08}$	$-1.840232 \times 10^{-08}$	$-2.234417 \times 10^{-08}$
$a_5$	$9.889105 \times 10^{-11}$	$1.133240 \times 10^{-10}$	$1.270510 \times 10^{-10}$

En la siguiente figura, se encuentran las gráficas de las tasas de mortalidad de la tabla "CNSF 2000 - I (1991-1998)" y de las tasas calculadas con los coeficientes antes obtenidos:



El error entre las tasas calculadas y las tasas de la tabla de mortalidad se encuentra representado en las siguientes dos gráficas. Se debe observar particularmente como el error crece rápidamente conforme se amplía el rango de edades a ajustar:





## 4.5 Alcances y Limitaciones.

### Alcances.

Gracias a su flexibilidad, el presente método permite escoger el modelo con el adecuado balance entre sencillez y exactitud. Cuando el grado del polinomio se escoja muy alto se tendrá una muy buena exactitud pero el modelo tendrá demasiados parámetros (términos del polinomio). Si por el contrario el grado del polinomio es demasiado pequeño, el resultado será un polinomio muy sencillo pero con grandes desviaciones respecto a los datos.

El manejo del error relativo en el modelo permite usar polinomios bastante sencillos que ajustan de manera aceptable a todos los datos de la tabla.

La función obtenida, un polinomio, es una función muy sencilla que se presta al manejo de muchas hipótesis y desarrollos de trabajo.

Permite, en contraste con el método del capítulo 3, ajustar tablas de mortalidad con un comportamiento errático o atípico en las cuales las tasas de mortalidad no se incrementan conforme se incrementa la edad. Un claro ejemplo es la tabla de mortalidad para hombres presentada en el trabajo "Tablas de Mortalidad 2000" realizado por la AMIS-AMAC, la cual presenta un mínimo local a la edad de 27 años.

### Limitaciones.

La principal limitación del presente método es que su uso requiere de paquetería de computadora que pueda multiplicar e invertir matrices, como es el caso de Excel® o Matlab®.

El método es más largo y laborioso que el presentado en el capítulo 3.

Se pierden demás aplicaciones y formulas desarrolladas en el cálculo actuarial basados en la ley de Makeham.

## Conclusiones.

Del presente trabajo podemos concluir que:

La constante competencia entre compañías aseguradoras deriva en la búsqueda de nuevos y diversos productos de Seguros que satisfagan cada una de las necesidades del mercado asegurador mexicano. Además los productos deben tener los costos más atractivos posibles para garantizar un adecuado volumen de ventas dentro del margen de utilidades requerido.

Los métodos de ajuste presentados en este trabajo, son un poderoso Instrumento para desarrollar productos de gran diversidad y bajo costo:

- Por un lado, cada método de ajuste resulta en un modelo matemático que puede emplearse en el cálculo de las más diversas formas de Seguros continuos de vida.
- Por el otro, los métodos permiten a las compañías aseguradoras ajustar continuamente las experiencias propias de mortalidad y ofrecer productos que pudieran ser más atractivos que los calculados con la tabla de mortalidad "CNSF 2000 - I (1991 - 1998)".

El método de ajuste tradicional es muy sencillo de aplicar y pese a la sencillez de su fórmula tiene una muy buena bondad de ajuste. Sin embargo, pierde considerable precisión cuando se utiliza en rangos de edades muy amplos o con singularidades. Por ejemplo, sería totalmente inapropiado para graduar las experiencias de mortalidad de 0 a 18 años para el cálculo de Seguros educacionales.

La principal desventaja del método tradicional se debe a su rigidez, pues este se basa en la primera Ley de Makeham que no permite incorporar nuevos parámetros al modelo en la búsqueda de una mejor bondad de ajuste. Este problema, se suma al hecho que el método no hace un manejo del error lo cual no permite al modelo adaptarse de la mejor manera posible a los datos.

No obstante el método propuesto (a base de polinomios) resulta mucho más complicado en su aplicación y requiere poderosa paquetería de cálculo, presenta mejores resultados en términos del error que el método tradicional. Además es un método que al permitir la inclusión de nuevos parámetros al

sistema (grado del polinomio) permite mejorar el error según las necesidades y objetivos del usuario.

Por su flexibilidad y manejo del error, el método propuesta resulta mejor para ajustar amplios rangos de edades y/o porciones de la tabla donde se presentan singularidades reales como sería una tabla de 0 a 18 años.

Es necesario aclarar que, ambos métodos funcionan sobre tasas singulares de mortalidad que fueron previamente aisladas del las demás causas de decremento mediante otras metodologías aquí no presentadas. Si se requiere ajustar las experiencias propias de alguna compañía, previamente será necesario encontrar las tasas singulares de mortalidad con base a la teoría de Tablas de Decremento Múltiple para posteriormente aplicar los métodos de ajuste aquí ejemplificados.

Finalmente es de comentar, que las continuamente nuevas y más poderosas herramientas de cálculo, hacen posible el desarrollo de mejores (aun que casi siempre más complicados) modelos sobre la fuerza de mortalidad  $\mu_x$  que es la base del cálculo para vida.

---

# Anexo.

## Probabilidad.

La principal herramienta de estudio y desarrollo del cálculo actuarial, es sin lugar a dudas la teoría de las probabilidades. Esta herramienta resulta útil no solo para el cálculo actuarial, si no además para el análisis de todo suceso o experimento que involucre un resultado azaroso. Un sencillo ejemplo es el lanzamiento de una moneda, cuyos posibles resultados es que la moneda caiga sol o caiga águila.

Dada la importancia del tema, a continuación se presenta un breve curso que contiene los elementos primordiales de la teoría de probabilidades basado en la traducción de las primeras páginas del texto "Introduction to the theory of statistics" de Alexander M. Mood, así como en las notas en clase del curso de Probabilidad I del profesor Miguel Ángel García de la Facultad de Ciencias de la UNAM.

La Teoría de las Probabilidades tiene sus comienzos formales en el siglo XVII con los juegos de azar; baraja, dados, ruleta, etc. En este tipo de juegos, si nos involucramos desde el punto de vista del jugador, el mayor alcance que la probabilidad puede tener es ayudarnos para tomar ciertas decisiones, si es que estas son posibles, para maximizar nuestros chances de ganar. Sin embargo en este caso la probabilidad no podrá ayudarnos para contestarnos si ganaremos o perderemos.

Por el otro lado, y es donde la probabilidad tiene su mayor utilidad es desde el punto de vista del que diseña tales juegos de azar y de las Casas de Juego. El diseñador de los juegos de azar debe usar los resultados de la probabilidad para hacer juegos que produzcan ganancias a largo plazo aun que nunca se pueda predecir un resultado en un solo juego o a corto plazo. Es decir, la incierta ganancia o pérdida de un jugador en el corto plazo, asegura la cierta ganancia de la Casa de Juego al largo plazo.

De manera análoga, la probabilidad le resulta de poca utilidad al dueño de un coche para saber si éste será robado o chocado. Sin embargo las probabilidades resultan sumamente útiles para las aseguradoras pues les permite

---

desarrollar tarifas atractivas de seguros de automóviles que garanticen un margen de utilidad con un suficiente número de asegurados.

El comentario anterior tiene como objetivo, aclarar que las leyes de probabilidad no sirven de "magos" para predecir que resultado se obtendrá al realizar de manera única un experimento azaroso, si no más bien, sirven para determinar con que incidencia aparecerá cada resultado al realizar el experimento, lo que se hará evidente al repetir el experimento un número suficiente de veces.

Con la finalidad de poder continuar el entendimiento del presente tema, se dará por el momento una idea intuitiva (más no definición) de lo que es la probabilidad. La probabilidad es una función que relaciona a cada posible resultado del experimento con un número real entre cero y uno. Tal número no se elige al azar, si no que es un indicador de las expectativas que tiene dicho resultado de aparecer al realizar un experimento. Por ejemplo, una probabilidad de cero nos da a entender que dicho resultado es imposible que ocurra, mientras una probabilidad de uno nos indica que dicho resultado es seguro que aparezca al realizar el experimento.

$$P[\text{resultado}] = p, \text{ con } 0 \leq p \leq 1$$

Cuando repetimos muchas veces el experimento, dicho indicador se asemejará a la frecuencia relativa con la que aparece el resultado en cuestión.

Actualmente existen 3 enfoques o métodos para encontrar dicho número y dependiendo de las características propias de cada problema elegiremos entre uno o la combinación de 2 de ellos: probabilidad clásica o "a priori", probabilidad frecuentista o "a posteriori" y probabilidad axiomática.

El enfoque clásico o a priori se basa en el análisis de las condiciones del experimento para mediante un razonamiento deductivo encontrar el número buscado. Por otro lado el enfoque frecuentista o posteriori se basa en la realización explícita del experimento un número significativo de veces para utilizar la frecuencia relativa observada como una aproximación de la probabilidad del resultado.

El enfoque axiomático, el más importante y poderoso de los tres, permite resolver problemas de una complejidad mayor a los que permiten el enfoque clásico o el enfoque frecuentista. Esto es posible ya que la probabilidad

axiomática cuenta con una vasta cantidad de definiciones, teoremas, etc. que permiten la descomposición de un determinado problema en otro/s más sencillo/s. Es muy común que mediante la probabilidad axiomática, se descomponga un problema de "segundo nivel" en problemas más sencillos de "primer nivel" que se resuelven mediante definiciones más elementales.

## I. Enfoque Clásico o Probabilidad a Priori.

Supóngase el problema de tener que calcular la probabilidad que tiene una moneda de salir sol al ser lanzada al aire. De manera inmediata, sabemos que el experimento solo tiene dos posibles resultados; que la moneda caiga sol o caiga águila. Si además suponemos que ambos resultados son igualmente probables, basándonos en la suposición de que la moneda es aceptablemente simétrica, llegamos a la rápida conclusión que tal probabilidad es igual a  $\frac{1}{2}$ .

Este razonamiento es el que da lugar a la definición de probabilidad clásica:

### **Definición Probabilidad Clásica.**

Si un experimento aleatorio tiene solo  $n$  resultados posibles, mutuamente excluyentes e igualmente probables, y si  $n_A$  de estos resultados tienen un atributo o característica común  $A$ , entonces la *probabilidad* del atributo  $A$  es el cociente  $n_A/n$ .

Observe que la definición anterior es muy transparente y sencilla, por lo que siempre que se pueda utilizar será nuestro método favorito para resolver algún problema. Basta con que se haga una cuidadosa enumeración de todos y cada uno de los resultados del experimento, así como otra enumeración de todos los casos que cumplen con la propiedad que estamos midiendo, y finalmente dividir ambos números.

### **Ejemplo 1.**

Con la definición de Probabilidad Clásica, calcular la probabilidad que tiene una moneda de salir sol al ser lanzada al aire.

Puesto que la moneda solo puede caer sol o águila tenemos un total de posibles resultados es dos. Además, solo de uno de estos dos resultados cumple con la propiedad de que la moneda caiga sol. Entonces tenemos que:

$$n = 2$$

$A = \{\text{La moneda cae cara sol para arriba}\}$

$$n_A = 1$$

$$\therefore P[A] = n_A / n = 1/2$$

### **Ejemplo 2.**

Con base a la definición de Probabilidad Clásica, calcular la probabilidad de sacar al azar una carta as de una Baraja.

En total una baraja tiene un total de 52 cartas, de las cuales solamente 4 son cartas as. Por lo tanto la probabilidad de sacar una carta as es:

$$n = 52$$

$A = \{\text{Sacar una carta as}\}$

$$n_A = 4$$

$$\therefore P[A] = n_A / n = 4/52 = 1/13.$$

- ***Ventajas de la probabilidad clásica.***

La probabilidad clásica nos resulta una poderosa herramienta para el cálculo de probabilidades, con la ventaja de que no es necesario siquiera realizar el experimento explícitamente. Basta saber sumar, así como el uso de un razonamiento deductivo y suponer ciertas condiciones ideales para obtener de manera muy sencilla el resultado buscado.

- ***Precauciones al usar la probabilidad clásica.***

Al resolver problemas usando la definición de Probabilidad clásica, frecuentemente se pierde de vista el hecho que los  $n$  resultados deben ser mutuamente excluyentes e igualmente probables como lo dice la definición.

Especial cuidado requieren estas dos condiciones para que los resultados que se obtengan sean correctos.

Para hacer gráfico el tipo de errores en los que se puede caer al no considerar resultados mutuamente excluyentes, tome como ejemplo el problema de calcular la probabilidad de sacar una carta de tipo espada o as de una baraja. Si de manera inocente, argumentamos que una baraja tiene 13 cartas tipo espada y 4 cartas tipo as y que en total la baraja tiene 52 cartas diferentes, obtenemos el resultado de tal probabilidad es igual a  $17/52$ . Este precipitado resultado tiene el error de no considerar el hecho de que una de las cartas espada es a su vez una carta as, es decir, 2 de los 17 resultados no son excluyentes sin no más bien son el mismo. Por lo tanto, ahora que es evidente que solo estamos hablando de 16 cartas diferentes (resultados excluyentes) de entre 52 sabemos que el cálculo correcto es  $16/52$ .

Por otro lado, se puede caer en el error de considerar resultados que no sean igualmente probables (o equiprobables). Esta situación suele presentarse cuando no se consideran cada uno de los posibles resultados de un experimento en forma detallada. A modo de ejemplo, calculemos la probabilidad que tienen dos monedas de caer cara sol si son lanzadas al aire. Entonces se pudieran enumerar los posibles resultados de este experimento de la siguiente manera:  $R_1 = \{\text{que las dos monedas caigan sol}\}$ ,  $R_2 = \{\text{que una moneda caiga sol y la otra águila}\}$  y  $R_3 = \{\text{que las dos monedas caigan águila}\}$ . Si ahora se aplica la definición de probabilidad clásica sobre la enumeración anterior, se obtiene que la probabilidad de que las dos monedas caigan sol es igual a  $1/3$  lo cual es incorrecto.

La falla del razonamiento anterior es consecuencia de haber utilizado una enumeración de resultados no equiprobables; mientras que  $R_1$  solo se puede presentar de una manera (cuando las dos monedas caen sol al mismo tiempo),  $R_2$  se puede dar de 2 diferentes maneras (sol en la primera moneda y águila en la segunda o águila en la primera moneda y sol en la segunda). Entonces tenemos que en realidad hay 4 diferentes y equiprobables resultados posibles:  $R_1 = \{\text{sol y sol}\}$ ,  $R_2 = \{\text{sol y águila}\}$ ,  $R_3 = \{\text{águila y sol}\}$  y  $R_4 = \{\text{águila y águila}\}$ . De esta manera se obtiene que la probabilidad buscada es igual a  $1/4$ .

Como moraleja del último ejemplo, nos queda que para utilizar adecuadamente la probabilidad clásica, se deben enumerar de forma detallada todos los posibles resultados del experimento, de tal manera que estos sean lo más elementales e indivisibles posible. Este proceder por lo general nos garantiza que trabajemos con resultados excluyentes y además equiprobables.

---

- **Limitaciones de la probabilidad clásica**

La primera limitante de la probabilidad clásica es consecuencia directa de su definición. El total de posibles resultados  $n$  debe ser un número finito.

Otra limitante es que la probabilidad clásica supone que el experimento se realiza bajo condiciones ideales, para asegurar así la equiprobabilidad de los diferentes resultados. Esta suposición no siempre podrá ser cierta o ser tomada como cierta. Por ejemplo, si lanzamos una moneda que es significativamente desbalanceada, podremos afirmar que la probabilidad que la moneda caiga Sol no es igual a la probabilidad que la moneda caiga águila en cuyo caso no podremos aplicar la definición de probabilidad clásica.

Sin embargo, la más importante desventaja es el limitado campo de acción. Por sí sola, la probabilidad clásica solo puede resolver problemas probabilísticos muy elementales. Por tal razón, no debe ser vista como un todo si no más bien como parte del aprendizaje para darse una idea de lo que debe ser una función de probabilidad.

## II. Enfoque Frecuentista o Probabilidad a Posteriori.

La Probabilidad Frecuentista consiste en la realización explícita de un experimento, para utilizar las frecuencias relativas observadas como aproximación de la probabilidad.

Todo cálculo de una probabilidad hace referencia a un atributo o característica de un experimento dado. Suponga entonces que dicho experimento se puede repetir cuantas veces se quiera bajo condiciones similares. Entonces si se repite el experimento un número suficiente de veces, se puede usar la frecuencia relativa del atributo en cuestión como una aproximación a la probabilidad de dicho atributo. Si resumimos obtenemos la siguiente definición:

### Definición.

Si una serie de observaciones o experimentos, son realizados bajo condiciones uniformes, entonces la probabilidad que ocurra el evento A se puede aproximar mediante la frecuencia relativa del evento A.

Donde:

$$\text{Frecuencia relativa (A)} = \frac{\text{Ocurrencias Observadas de A}}{\text{No. de Veces que se repitió el Experimento}}$$

A fines prácticos, por el momento se podrá entender por *evento* como uno o más resultados del experimento, que tienen un atributo o característica en común.

### Ejemplo 3.

En la siguiente tabla se registraron las ocurrencias observadas al lanzar una moneda 100 veces al aire:

Evento	Ocurrencias registradas	Frecuencias Relativas Observadas	Frecuencias Relativas Esperadas
Sol	56	.56	.5
Águila	44	.44	.5
Total	100	1.0	1.0

Fuente : "Introduction to the theory of Statistics" de Mood Graybill Boes.

Utilizando la información de la tabla anterior, tenemos que si  $A = \{\text{La moneda cae sol}\}$  entonces la  $P[A] \approx .56$ .

De manera similar, en la siguiente tabla se encuentra el resumen de 300 lanzamientos de un dado bajo condiciones similares:

Evento	Ocurrencias registradas	Frecuencias Relativas Observadas	Frecuencias Relativas Esperadas
Cae 1 arriba	51	.170	.166
Cae 2 arriba	54	.180	.166
Cae 3 arriba	48	.160	.166
Cae 4 arriba	51	.170	.166
Cae 5 arriba	49	.163	.166
Cae 6 arriba	47	.157	.166
Total	300	1.000	1.000

Fuente : "Introduction to the theory of Statistics" de Mood Graybill Boes.

Observe que en las tablas anteriores, las frecuencias relativas observadas y las esperadas son similares, de manera que realmente es indiferente cual se utilice para el cálculo de una probabilidad.

- **Ventajas de la probabilidad frecuentista.**

No obstante existe un dominio común entre ambos métodos, son ventajas de la probabilidad frecuentista aquellos casos o problemas que la probabilidad clásica no pueda resolver.

Tómese como ejemplo el caso de la moneda desbalaceada. De igual manera, el calcular la probabilidad que tiene una persona de edad  $x$  de morir antes de tener la edad  $x+1$ , solo puede ser calculada utilizando la probabilidad frecuentista. Es decir, las tablas de mortalidad son posibles gracias al enfoque frecuentista o a posteriori en complemento de otros métodos.

---

- **Limitaciones de la probabilidad frecuentista.**

Al igual que la probabilidad clásica, la principal desventaja del enfoque frecuentista es que, por sí solo, es muy reducido el universo de problemas que puede resolver.

Otra limitante es que es un método muy poco práctico. Para resolver problemas realmente elementales es necesario repetir múltiples veces un experimento, para obtener frecuencias relativas confiables.

### III. Probabilidad Axiomática.

Uno de los principales objetivos de la ciencia, es poder predecir o describir hechos o fenómenos del entorno en que vivimos. Un camino idóneo para lograrlo, es mediante el desarrollo de modelos matemáticos.

En la probabilidad Axiomática desarrollaremos modelos matemáticos que puedan ser usados para describir eventos del mundo real. Modelos a problemas muy elementales pueden ser desarrollados utilizando la probabilidad clásica o la probabilidad frecuentista. Sin embargo, problemas de complejidad mayor sólo podrán ser resueltos con el apoyo de definiciones, procedimientos, axiomas y propiedades de los cuales consta la probabilidad axiomática.

Para empezar, requerimos que todo posible resultado de un experimento bajo estudio pueda ser numerado. Por ejemplo, el lanzamiento de una moneda involucra solamente dos resultados posibles; sol o águila. Asociaremos probabilidades solamente a este tipo de resultados o colecciones de resultados.

Incluso si un resultado es imposible que ocurra, este puede incluirse (su probabilidad es cero). Lo que es importante de recordar, es que todo resultado que pueda ocurrir debe ser incluido.

De ahora en adelante, todo posible resultado del experimento bajo estudio debe ser visto como un punto muestral y al conjunto de todos los posibles resultados o puntos muestrales se define como espacio muestral. Este manejo del experimento como conjuntos y elementos hace necesario el dominio de la teoría de conjuntos.

### 3.1 Teoría de Conjuntos.

Comenzamos con una colección de objetos. Cada objeto en nuestra colección será llamado *punto* o *elemento*. Asumimos que nuestra colección de objetos es suficientemente grande para incluir todos los puntos bajo consideración en una discusión dada. El conjunto de todos estos puntos es llamado *espacio*, *universo* o *conjunto universal* y será denotado por medio de una  $\Omega$ . Sea denotará por  $\omega$  a los elementos o puntos en  $\Omega$ . Aunque un conjunto puede ser definido como una colección de cualquier tipo de objetos, asumiremos a menos que se indique lo contrario, que todo conjunto consiste solamente de puntos en el espacio  $\Omega$ .

#### Ejemplo 4.

$\Omega = \mathbb{R}_2$ , donde  $\mathbb{R}_2$  es el conjunto de puntos  $\omega$  en el plano  $\{\omega = (x, y) \mid x \text{ y } y \in \mathbb{R}\}$ .

#### Ejemplo 5.

$\Omega = \{\text{Todos los ciudadanos de los Estados Unidos}\}$ .

Por lo general usaremos las primeras letras del alfabeto latín en mayúsculas, con o sin subíndices, para denotar conjuntos. Si  $\omega$  es un punto o un elemento que pertenece al conjunto  $A$ , {escribiremos que, lo denotaremos como}  $\omega \in A$ . Si  $\omega$  no es un elemento del conjunto  $A$ , {escribiremos que, lo denotaremos como}  $\omega \notin A$ .

#### Definición 1, Subconjunto.

Si todo elemento del, conjunto  $A$  es también elemento del conjunto  $B$ , entonces se dice que  $A$  es un *subconjunto* de  $B$ , y lo denotaremos como  $A \subset B$  ó  $B \supset A$ .  $A \subset B$ ; se debe leer como "A esta contenido en B" y  $A \subset B$  como "B contiene a A".

#### Definición 2, Conjuntos equivalentes.

Se dice que dos conjuntos son *equivalentes* o *iguales*, si  $A \subset B$  y  $B \subset A$ . Lo escribiremos como  $A = B$ .

### Definición 3, Conjunto Vacío.

Si el conjunto  $A$  no contiene ningún elemento, será llamado *conjunto vacío* o *conjunto nulo* y se denotará por  $\emptyset$ .

### Definición 4, Complemento.

El complemento de un conjunto  $A$  con respecto al espacio  $\Omega$ , denotado por  $\bar{A}$ ,  $A^c$  ó  $\Omega - A$ , es el conjunto de todos los puntos que están en  $\Omega$  pero que no están en  $A$ .

### Definición 5, Unión.

Sean  $A$  y  $B$  dos conjuntos cualquiera del espacio  $\Omega$ , el conjunto formado por todos los puntos que están en  $A$  o en  $B$  o en ambos es definido como la *unión* de  $A$  y  $B$  y se escribe  $A \cup B$ . Por convención se lee como "A unión B".

### Definición 6, Intersección.

Sean  $A$  y  $B$  dos conjuntos cualquiera del espacio  $\Omega$ , el conjunto formado por todos los puntos que están en  $A$  y que también están en  $B$  es definido como la *intersección* de  $A$  y  $B$  y se escribe  $A \cap B$  o simplemente  $AB$ . Por convención se lee como "A intersección B".

### Definición 7, Diferencia.

Sean  $A$  y  $B$  dos conjuntos cualesquiera del espacio  $\Omega$ . Al conjunto de todos los puntos que están en  $A$  y no en  $B$  se define como el conjunto *diferencia* de  $A$  y  $B$ ; se denota por  $A - B$  y por convención se lee como "A menos B".

### Ejemplo 6.

Sea  $\Omega = \{(x, y) \mid 0 \leq x \leq 1 \text{ y } 0 \leq y \leq 1\}$ , lo cual se lee "sea  $\Omega$  el espacio formado por la colección de todos los puntos  $(x, y)$  para los cuales  $0 \leq x \leq 1$  y  $0 \leq y \leq 1$ ", y sean los siguientes conjuntos:

$$A_1 = \{(x, y) \mid 0 \leq x \leq 1 \text{ y } 0 \leq y \leq \frac{1}{2}\}$$

$$A_2 = \{(x, y) \mid 0 \leq x \leq \frac{1}{2} \text{ y } 0 \leq y \leq 1\}$$

$$A_3 = \{(x, y) \mid 0 \leq x \leq y \leq 1\}$$

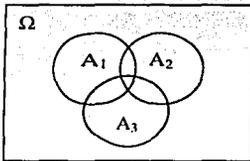
$$A_4 = \{(x, y) \mid 0 \leq x \leq \frac{1}{2} \text{ y } 0 \leq y \leq \frac{1}{2}\},$$

se tienen las siguientes relaciones:

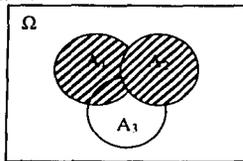
$$\begin{aligned}
 A_4 &\subset A_1; \\
 A_4 &\subset A_2; \\
 A_1 \cap A_2 &= A_4; \\
 A_2 \cup A_3 &= A_4 \cup A_3; \\
 \bar{A}_1 &= A_1^c = \{(x, y) \mid 0 \leq x \leq 1 \text{ y } \frac{1}{2} \leq y \leq 1\} \\
 A_1 - A_4 &= \{(x, y) \mid \frac{1}{2} \leq x \leq 1 \text{ y } 0 \leq y \leq \frac{1}{2}\}
 \end{aligned}$$

Nótese que utilizamos llaves para reunir o englobar a los puntos de un conjunto.

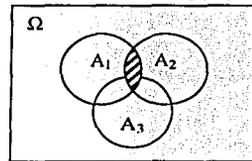
En los siguientes diagramas de Venn se muestran  $\Omega$ ,  $A_1$ ,  $A_2$  y  $A_3$ , así como las operaciones entre los mismos.



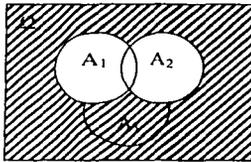
$\Omega$ ,  $A_1$ ,  $A_2$  y  $A_3$



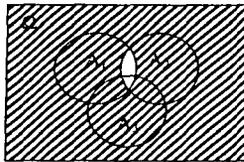
$A_1 \cup A_2$



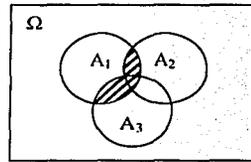
$A_1 \cap A_2$



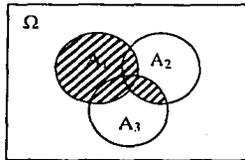
$(A_1 \cup A_2)^c = A_1^c \cap A_2^c$



$(A_1 \cap A_2)^c = A_1^c \cup A_2^c$



$A_1 \cap (A_1 \cup A_2) = (A_1 \cap A_2) \cup (A_1 \cap A_3)$



$A_1 \cup (A_2 \cap A_3) = (A_1 \cup A_2) \cap (A_1 \cup A_3)$

TESIS CON  
FALLA DE ORIGEN

Las operaciones de unión, intersección y complemento satisfacen un gran número de leyes, algunas de las cuales se citan a continuación han sido aceptadas como teoremas. La demostración se omite.

### **Teorema 1, Leyes Conmutativas.**

$$A \cup B = B \cup A$$

$$A \cap B = B \cap A$$

### **Teorema 2, Leyes Asociativas.**

$$A \cup (B \cup C) = (A \cup B) \cup C$$

$$A \cap (B \cap C) = (A \cap B) \cap C$$

### **Teorema 3, Leyes distributivas.**

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

### **Teorema 4.**

$(A^c)^c = A$ , es decir, el complemento del complemento de A es igual a A.

### **Teorema 5.**

$$A \cap \Omega = A$$

$$A \cup \Omega = \Omega$$

$$A \cap \emptyset = \emptyset$$

$$A \cup \emptyset = A$$

### **Teorema 6.**

$$A \cap A^c = \emptyset$$

$$A \cup A^c = \Omega$$

$$A \cap A^c = A$$

$$A \cup A^c = A$$

### **Teorema 7, Leyes de Morgan.**

$$(A \cup B)^c = A^c \cap B^c$$

$$(A \cap B)^c = A^c \cup B^c$$

**Teorema 8.**

$$A - B = A \cap B^c$$

De manera ilustrativa, se demostrará una de las leyes de Morgan.

Para demostrar la igualdad entre dos conjuntos, lo primero que se debe intentar es aplicar la definición de igualdad entre conjuntos, es decir, demostrar la doble contención.

Caso 1. Por demostrar que  $(A \cup B)^c \subset A^c \cap B^c$ . Nos referimos a la definición de contención:

$$\begin{aligned} \text{Si } \omega \in (A \cup B)^c \text{ tenemos que } \omega \notin (A \cup B) \\ \Rightarrow \omega \notin A \text{ y } \omega \notin B \\ \Rightarrow \omega \in A^c \text{ y } \omega \in B^c \\ \Rightarrow \omega \in (A^c \cap B^c) \\ \therefore (A \cup B)^c \subset (A^c \cap B^c) \end{aligned}$$

Caso 2. Por demostrar que  $A^c \cap B^c \subset (A \cup B)^c$ .

$$\begin{aligned} \text{Si } \omega \in A^c \cap B^c \text{ tenemos que } \omega \in A^c \text{ y que } \omega \in B^c \\ \Rightarrow \omega \notin A \cup B, \text{ por que de lo contrario, se tendría que } \omega \text{ al menos pertenece} \\ \text{a alguno de } A \text{ o } B, \text{ contradiciendo el hecho de que } \omega \in A^c \text{ y } B^c \\ \Rightarrow \omega \in (A \cup B)^c \\ \therefore (A^c \cap B^c) \subset (A \cup B)^c \end{aligned}$$

Entonces, por cumplirse la doble contención se tiene que:

$$(A \cup B)^c = (A^c \cap B^c) \quad \blacksquare$$

La unión e intersección fueron definidas únicamente sobre un par conjuntos. Tales definiciones se pueden extender para más de 2 conjuntos, de hecho para un número arbitrario de conjuntos.

Es costumbre distinguir los diferentes conjuntos de entre una colección de subconjuntos de  $\Omega$ , asignándoles diferentes subíndices. Así sabremos que dos conjuntos son diferentes pero pertenecen a una misma colección, por tener mismo nombre pero diferentes subíndices.  $\Lambda$  (letra lambda mayúscula del alfabeto griego) denota un catálogo de nombres o índices.  $\Lambda$  es llamada también *conjunto índice*. Por ejemplo, si nosotros estamos interesados únicamente con 2 conjuntos, entonces nuestro conjunto índice  $\Lambda$  incluye solamente 2 índices;  $\Lambda = \{1, 2\}$ .

### Definición 8, Unión e intersección de conjuntos.

Sea  $\Lambda$  un conjunto índice y sea  $\{A_\lambda \mid \lambda \in \Lambda\} = \{A_\lambda\}$  una colección de subconjuntos de  $\Omega$  ordenados por  $\Lambda$ . El conjunto de puntos que consta de todos los puntos que pertenecen a  $A_\lambda$  para al menos un  $\lambda$  en  $\Lambda$ , es llamado *unión* de los conjuntos  $\{A_\lambda\}$  y se denota por  $\bigcup_{\lambda \in \Lambda} A_\lambda$ . El conjunto de puntos que consta de todos los puntos que pertenecen a cada  $A_\lambda$  para todo  $\lambda$  en  $\Lambda$ , es llamado *intersección* de los conjuntos  $\{A_\lambda\}$  y se denota por  $\bigcap_{\lambda \in \Lambda} A_\lambda$ .

#### Ejemplo 7.

Si  $\Lambda = \{1, 2, \dots, N\}$ , es decir,  $\Lambda$  es el conjunto índice que consiste de los primeros  $N$  enteros, entonces  $\bigcup_{\lambda \in \Lambda} A_\lambda$  también se puede escribir como:

$$\bigcup_{n=1}^N A_n = A_1 \cup A_2 \cup \dots \cup A_N.$$

Uno de los teoremas más fundamentales que relacionan uniones, intersecciones y complementos para una colección arbitraria de conjuntos es justamente el Teorema *De Morgan*.

#### Teorema 9, Teorema de Morgan.

Sea  $\Lambda$  un conjunto índice y  $\{A_\lambda\}$  una colección de subconjuntos de  $\Omega$  ordenada por  $\Lambda$ . Entonces,

$$(i) \left( \bigcup_{\lambda \in \Lambda} A_\lambda \right)^c = \bigcap_{\lambda \in \Lambda} A_\lambda^c$$

$$(ii) \left( \bigcap_{\lambda \in \Lambda} A_\lambda \right)^c = \bigcup_{\lambda \in \Lambda} A_\lambda^c$$

No se dará una demostración del Teorema de Morgan, sin embargo, nótese que éste es la generalidad del Teorema 7 donde  $\Lambda = \{1, 2\}$  cuya demostración se realizó a medias previamente. Es prudente hacer notar además, que si  $\Lambda$  es definido a partir de los números naturales, entonces el Teorema 9 se puede demostrar fácilmente utilizando inducción matemática sobre el Teorema 7.

### **Definición 9, Conjuntos disconexos o mutuamente excluyentes.**

Los subconjuntos A y B de  $\Omega$  se define que son mutuamente excluyentes si  $A \cap B = \emptyset$

La colección de subconjuntos  $\{A_i\}$  se dice que son mutuamente excluyentes si se tiene que  $A_i \cap A_j = \emptyset \forall i, j \in \Lambda$ .

#### **Teorema 10.**

Si A y B son subconjuntos de  $\Omega$ , se tiene que,

$$(i) A = (A \cap B) \cup (A \cap B^c)$$

$$(ii) (A \cap B) \cap (A \cap B^c) = \emptyset$$

#### **Demostración.**

$$(i) (A \cap B) \cup (A \cap B^c) = A \cap (B \cup B^c) = A \cap (\Omega) = A$$

$$(ii) (A \cap B) \cap (A \cap B^c) = A \cap B \cap A \cap B^c = (A \cap A) \cap (B \cap B^c) = A \cap \emptyset = \emptyset \quad \blacksquare$$

#### **Teorema 11.**

Si  $A \subset B$ , entonces,

$$(i) A \cap B = A$$

$$(ii) A \cup B = B$$

#### **Demostración**

La demostración se deja como ejercicio. Recuerde el lector para demostrar la igualdad entre conjuntos lo mejor intentar aplicar directamente la definición, es decir, demostrar la doble contención.

### **3.2 Definición de Espacio muestral y evento.**

Hemos descrito con anterioridad lo que podría entenderse por modelo de probabilidad. Tenemos en mente de manera conceptual algunos experimentos cuyos posibles resultados nos gustaría estudiar, cuantificando la probabilidad de ciertos resultados o colección de resultados. En esta sección, se darán dos importantes definiciones, junto con algunos ejemplos, que serán usados para cuantificar tales probabilidades

### **Definición 10, Espacio Muestral.**

El *espacio muestral*, denotado por  $\Omega$ , es la colección de todos los posibles resultados de un experimento conceptual. Se entenderá por *punto o elemento del espacio muestral* a cualquier posible resultado del experimento.

Se podría entender la definición observando de manera individual las palabras que la componen. El uso de la palabra "espacio" se puede justificar ya que el espacio muestral es la totalidad de objetos o elementos, los cuales en este caso son resultados posibles del experimento. Esto es, se está cuidando el significado de Teoría de Conjuntos de la palabra "espacio", como todos los objetos de interés en una discusión dada. La palabra "muestral" se refiere al hecho de que nuestro experimento es aleatorio, lo que significa que sus posibles resultados son inciertos, por lo que un resultado obtenido no es otra cosa que una muestra de entre varios posibles resultados.

### **Definición 11 , Evento y Espacio de Eventos.**

El *evento* es un subconjunto de un espacio muestral. La clase de todos los eventos asociados con un experimento dado es definida *espacio de eventos*.

Lo anterior no define con precisión que es un evento. Un evento siempre será un subconjunto del Espacio Muestral, sin embargo, no todo subconjunto del Espacio Muestral es un Evento. Serán eventos aquellos subconjuntos del espacio muestral para los cuales exista asociada alguna probabilidad de la cual se pueda hablar al respecto.

De igual manera la clase de todos los subconjuntos (no eventos) del espacio muestral no necesariamente corresponderá con el espacio de eventos. Sin embargo, veremos que el espacio de eventos siempre puede ser seleccionado de tal manera que sea lo suficientemente grande para incluir todos aquellos subconjuntos que sean eventos.

Si un espacio muestral consta de solamente un número finito de puntos o de un número infinito pero numerable, entonces el correspondiente espacio de eventos y la clase de todos los subconjuntos del espacio muestral serán iguales. Dicho de otra manera, si el espacio muestral consta de un conjunto *a lo más numerable* de puntos, entonces todo subconjunto del espacio muestral será a su vez un evento del espacio muestral. Se dice que un conjunto es *a lo más numerable* si es finito o si es infinito pero numerable.

Nuestro interés principal no son los eventos por sí solos, si no más bien la probabilidad que tiene un evento de ocurrir o no ocurrir. Se dice que un evento  $A$  ocurre, si existe al menos un posible resultado del experimento que pertenece a  $A$ . Ya que todo punto, digamos  $\omega$ , es un subconjunto (un subconjunto que consiste de un solo punto  $\omega$ ) del espacio muestral  $\Omega$ , es candidato para ser un evento. De esta manera  $\omega$  puede ser visto como un punto en  $\Omega$  o como un subconjunto de  $\Omega$ . Para distinguirlo, escribiremos  $\{\omega\}$  en vez de simplemente  $\omega$  aun que de cualquier manera  $\omega$  siempre puede ser visto como un subconjunto de  $\Omega$ . Tales subconjuntos de un solo punto serán siempre eventos y se llamarán *eventos elementales*. También  $\emptyset$  y  $\Omega$  son subconjuntos de  $\Omega$  y también son eventos.  $\Omega$  es usualmente llamado *evento seguro* y  $\emptyset$  *conjunto vacía*.

Con excepción de  $\emptyset$  y  $\Omega$  procuraremos usar solamente las primeras Letras del alfabeto latín en mayúsculas, con o sin acentos, para denotar eventos. El espacio de eventos siempre lo denotaremos con letras mayúsculas en manuscritas del alfabeto latín, que generalmente serán  $\mathcal{A}$  y  $\mathcal{B}$ .

El espacio muestral es fundamental y por lo general es fácil de definir para un experimento dado. Sin embargo, como veremos, lo que realmente es esencial para definir una probabilidad es determinar el espacio de eventos. A continuación unos ejemplos de cada uno para irnos familiarizando.

### Ejemplo 8.

El experimento consiste en lanzar un dado regular de 6 caras. El dado podrá caer solamente con alguna de las 6 caras hacia arriba, por lo tanto hay 6 posibles resultados del experimento:

$$\Omega = \left\{ \begin{array}{|c|c|} \hline \cdot & \cdot \\ \hline \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|} \hline \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|} \hline \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|} \hline \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|} \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|} \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} \right\}$$

Ahora revisemos como es nuestro espacio de eventos  $\mathcal{A}$  para este  $\Omega$ . Tómese en cuenta que por tratarse de un experimento con espacio muestral finito,  $\mathcal{A}$  estará dado por todos los posibles subconjuntos de  $\Omega$ . Veamos los siguientes eventos:  $A = \{\text{cae cara numero par}\} = \left\{ \begin{array}{|c|c|} \hline \cdot & \cdot \\ \hline \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|} \hline \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|} \hline \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} \right\}$ .  $B = \{\text{cae cara numero non}\} = \left\{ \begin{array}{|c|c|c|} \hline \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|c|} \hline \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} , \begin{array}{|c|c|c|} \hline \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} \right\}$ .  $A_i = \{\text{cae cara } i\}$  con  $i = 1, 2, 3, 4, 5$  o  $6$ , es decir,  $A_1 = \left\{ \begin{array}{|c|} \hline \cdot \\ \hline \cdot \\ \hline \end{array} \right\}$ ,  $A_2 = \left\{ \begin{array}{|c|} \hline \cdot & \cdot \\ \hline \cdot & \cdot \\ \hline \end{array} \right\}$ , ...,  $A_6 = \left\{ \begin{array}{|c|} \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \end{array} \right\}$ . Así podemos seguir agregando elementos a nuestra clase de eventos, de tal manera que cuando se tengan incluidos todos, habremos armado nuestro  $\mathcal{A}$ . En éste caso,  $\mathcal{A}$  consta de  $2^6=64$  eventos diferentes, de los cuales solo 6 ( $A_1, A_2, A_3, A_4, A_5$  y  $A_6$ ) son eventos elementales.

En general para un espacio muestral finito de tamaño  $n$ , se tiene que esta formada por un total de  $2^n$  diferentes eventos.

### Ejemplo 9.

Se lanzan al aire una moneda de 10 centavos, una de 50 centavos y una moneda de 1 peso y se observa en cada una cual cara queda arriba. Existen entonces 8 posibles resultados en el experimento.

$$\Omega = \{(S, S, S), (S, S, A), (S, A, S), (S, A, A), \\ (A, S, S), (A, S, A), (A, A, S), (A, A, A)\}$$

Al arreglo  $(\cdot, \cdot, \cdot)$  se le conoce como tupla o  $n$ -ada, y se está reservando la primera posición para registrar el resultado de la moneda de 10 centavos, la segunda posición para el resultado de la moneda de 50 centavos y la tercera posición para la moneda de 1 peso. Sea  $A_i = \{\text{caen exactamente } i \text{ caras sol}\}$  con  $i = 0, 1, 2, \text{ o } 3$ , es decir,  $A_0 = \{(A, A, A)\}$ ,  $A_1 = \{(S, A, A), (A, S, A), (A, A, S)\}$ ,  $A_2 = \{(S, S, A), (S, A, S), (A, S, S)\}$ ,  $A_3 = \{(S, S, S)\}$ ; para toda  $i$ , se tiene que  $A_i$  es un evento. Nótese que solamente  $A_0$  y  $A_3$  son eventos elementales. De nuevo se tiene que todo subconjunto de  $\Omega$  es un evento y en total hay  $2^8 = 256$ .

### Ejemplo 10.

El experimento consiste en registrar el número de muertes en accidentes de tránsito en el estado de Nuevo León que habrá el siguiente año. Solo se puede tener como resultado un número entero no negativo. Entonces  $\Omega = \{0, 1, 2, \dots\}$ . Sea  $A = \{\text{menos de 500 muertes}\} = \{0, 1, 2, \dots, 499\}$ .  $A_i = \{\text{exactamente ocurran } i \text{ muertes}\}$  con  $i = 0, 1, 2, \dots$ , es decir,  $A_i = \{i\}$ . Por lo tanto éste espacio muestral consta de un número infinito de puntos, cada uno de los cuales a su vez es un evento elemental  $(A_i)$ .

Este es el caso de un espacio muestral infinito, pero ya que sus elementos son numerables, todo subconjunto del mismo es a su vez un evento.

### Ejemplo 11.

Seleccionar un foco y registrar el número exacto de horas que dura encendido antes de fundirse. Cualquier número real no negativo puede ser resultado del experimento. Entonces  $\Omega = \{x \mid x \geq 0\}$ . Por lo tanto para este espacio muestral se tiene un número infinito y no numerable de puntos por lo que no todo subconjunto de  $\Omega$  es a su vez un evento. Sin embargo, cualquier

subconjunto que se pueda probar será a su vez un evento. Por ejemplo,  $A = \{\text{el foco se funde después de } n \text{ y antes de } m \text{ horas}\} = \{x \mid n \leq x \leq m\}$  es un evento. Por el contrario, el subconjunto  $B = \{\text{que se funda exactamente en } 1000.00 \text{ horas}\}$  de  $\Omega$  no es un evento.

### Ejemplo 12.

Considere el experimento aleatorio que consiste por un lado de registrar el número de veces que llueve y por otro el total de centímetros de agua que llueve, durante el siguiente mes de Julio en el Distrito Federal. El espacio muestral podría representarse de la siguiente manera:

$$\Omega = \{(i, x) \mid i = 0, 1, 2, \dots, 0 \leq x\},$$

en donde la primera posición de la tupla indica el número de veces que llovió y la segunda posición indica el total de centímetros que llovió. Sea  $\omega = (7, 2.251)$  es un punto del espacio muestral  $\Omega$ , sin embargo, no es un evento. Mientras tanto,  $A = \{(7, x) \mid 2.0 \leq x \leq 2.5\}$  si es un evento.

### Ejemplo 13.

En un experimento, se examinó la producción de cinco variedades de trigo. Las cinco variedades fueron plantadas y crecen bajo condiciones similares.  $\Omega = \{(y_1, y_2, y_3, y_4, y_5) \mid y_i \geq 0, i = 1, 2, 3, 4, 5\}$  donde  $y_i$  representa la producción en litros por acre de la  $i$ -ésima variedad de trigo. Sea  $A$  el evento definido por la condición de que  $y_2, y_3, y_4$  y  $y_5$  se cada una al menos 0 litros mayor que  $y_1$ , la variedad estándar. Al evento  $A$  lo podemos escribir:

$$A = \{(y_1, y_2, y_3, y_4, y_5) \mid y_j \geq y_1 \geq 0, j = 2, 3, 4, 5\}.$$

Hasta este momento, nuestra definición de espacio muestral es precisa y completamente satisfactoria, más no así las definiciones de evento y espacio de eventos. Dijimos que el espacio de eventos debía ser suficiente grande para contener cualquier evento del experimento y que no todo subconjunto del espacio muestral podía ser un evento. Sin embargo, no hemos dicho con exactitud cuales subconjuntos del espacio muestral pueden ser eventos y cuales no pueden. Puesto que todavía no contamos con los elementos matemáticos suficientes para una definición formal, permítanos por el momento enunciar algunas propiedades que los eventos y el espacio de eventos deben cumplir:

- (i)  $\Omega \in \mathcal{A}$   
 (ii) Si  $A \in \mathcal{A}$ , entonces  $A^c \in \mathcal{A}$   
 (iii) Si  $A_1$  y  $A_2 \in \mathcal{A}$ , entonces  $A_1 \cup A_2 \in \mathcal{A}$ .

Hemos dicho que los eventos por si solos no son nuestro interés principal, si no más bien la probabilidad que tiene el evento de ocurrir. Seguramente para tal propósito necesitaremos que el evento seguro  $\Omega$  este incluido en el espacio de eventos. Además, si  $A$  debe ser un evento para que podamos hablar de la probabilidad de que éste ocurra, entonces  $A^c$  debe ser también evento para que podamos hablar de la probabilidad de que  $A$  no ocurra. De manera similar, si  $A_1$  y  $A_2$  son eventos, lo debe ser también el conjunto  $A_1 \cup A_2$ .

Una colección de eventos que cumplen con las propiedades (i), (ii) y (iii) es llamado *álgebra Booleana* de eventos o simplemente *álgebra* de eventos. Debemos aclarar que también la colección de todos los subconjuntos de  $\Omega$  satisface las propiedades (i), (ii) y (iii). Los siguientes teoremas se desprenden directamente de las propiedades de  $\mathcal{A}$ .

### Teorema 12.

$$\emptyset \in \mathcal{A}$$

#### Demostración.

$$\begin{aligned} \Omega &\in \mathcal{A} && \text{por (i)} \\ \Rightarrow \Omega^c &\in \mathcal{A} && \text{por (ii)} \\ \Rightarrow \Omega^c = \emptyset &\in \mathcal{A} && \blacksquare \end{aligned}$$

### Teorema 13.

Si  $A_1$  y  $A_2 \in \mathcal{A}$ , entonces  $A_1 \cap A_2 \in \mathcal{A}$ .

#### Demostración.

$$\begin{aligned} A_1 \cup A_2 &\in \mathcal{A} && \text{por (iii)} \\ \Rightarrow (A_1 \cup A_2)^c &\in \mathcal{A} && \text{por (ii)} \\ \Rightarrow (A_1^c \cap A_2^c) &\in \mathcal{A} && \text{por Leyes de Morgan} \\ \Rightarrow ((A_1^c)^c \cap (A_2^c)^c) &\in \mathcal{A} && \text{por (ii)} \\ \Rightarrow (A_1 \cap A_2) &\in \mathcal{A} && \text{por Teorema 4} \quad \blacksquare \end{aligned}$$

**Teorema 14.**

Si  $A_1, A_2, \dots, A_n \in \mathcal{A}$ , entonces  $\bigcup_{i=1}^n A_i$  y  $\bigcap_{i=1}^n A_i \in \mathcal{A}$ .

**Demostración.**

Por inducción matemática.

Puesto que todo espacio de eventos  $\mathcal{A}$  es a su vez un álgebra, se justifica el uso del símbolo  $\mathcal{A}$  (notación para una álgebra) para denotarlo.

Sin embargo hasta aquí, no hemos explicado por que la colección de subconjuntos de  $\Omega$  no puede siempre ser tomada por el espacio de eventos  $\mathcal{A}$ . Tal explicación se dará en la siguiente sección, una vez que se halla definido formalmente a la función de probabilidad.

**3.3 Definición de Probabilidad.**

En esta sección se dará la definición axiomática de probabilidad. Aunque esta definición formal por si sola no es suficiente para asignar probabilidades a eventos, si es un importante paso para lograrlo. Recuerde que el calcular probabilidades a resultados de un experimento es equivalente a calcular probabilidades sobre los eventos que relacionan tales resultados. Ya que la probabilidad, así como otros conceptos venideros serán definidos como funciones, empezamos esta sección repasando la definición de función.

**Definición típica de función.**

Una *función*, digamos  $f(\cdot)$ , es una regla (ley, fórmula o receta) que asocia a cada elemento de un conjunto de puntos con un y solo un elemento de otro conjunto de puntos. El primer conjunto de puntos, digamos  $A$ , es llamado *dominio* y el segundo conjunto es llamado *contradominio*.

**Definición 12, Función.**

Una *función*, digamos  $f(\cdot)$ , con dominio  $A$  y contradominio  $B$ , es una colección de pares ordenados, digamos  $(a, b)$  que satisfacen las siguientes condiciones:

- (i)  $a \in A$  y  $b \in B$ ,

- (ii) Para todo  $a \in A$  siempre existe un par ordenado  $(a, b)$  que lo contiene (aun que no necesariamente todo  $b \in B$  se encuentra en alguna pareja ordenada)
- (iii) No existan pares ordenados  $(a, b)$  distintos en los que se repita el primer elemento.

Si  $(a, b) \in f(\cdot)$ , escribiremos que  $f(a)=b$  y lo leeremos como "f de a es igual a b". A  $f(a)$  lo llamaremos el *valor* de  $f(\cdot)$  en  $a$  o *imagen* de  $a$  bajo  $f(\cdot)$ . Como ya se dijo, pueden existir puntos del contradominio que no tengan asociada ninguna pareja ordenada de la función. El *rango* de la función  $f(\cdot)$  es precisamente el conjunto de puntos del contradominio que si tienen asociada una pareja ordenada de la función: el *rango* de  $f(\cdot) = \{b \in B \mid f(a) = b \text{ para algún } a \in A\}$ . Se tiene que el rango de la función es un subconjunto del contradominio que puede o no ser igual a esta.

### Ejemplo 14.

Sean  $f_1(\cdot)$  y  $f_2(\cdot)$  dos funciones que tienen el conjunto de los números reales por dominio y contradominio, definidas por

$$f_1(\cdot) = \{(x, y) \mid y = x^3 + x + 1, -\infty < x < \infty\}, \text{ y}$$

$$f_2(\cdot) = \{(x, y) \mid y = x^2, -\infty < x < \infty\}.$$

El rango y contradominio de  $f_1(\cdot)$  son iguales, es decir todos los números reales. Sin embargo, el rango de  $f_2(\cdot)$  es el conjunto de los números reales positivos que no es igual al contradominio.

De particular interés para nosotros será un tipo particular de función llamado función indicadora:

### Definición 13, Función Indicadora.

Sea  $\Omega$  cualquier espacio muestral con puntos  $\omega$  y  $A$  cualquier subconjunto de  $\Omega$ . La función indicadora de  $A$ , denotada por  $I_A(\cdot)$ , es la función con dominio  $\Omega$  y contradominio  $= \{0, 1\}$  definida por

$$I_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

Claramente se aprecia que la función  $I_A(\cdot)$  indica o describe al conjunto  $A$ .

## Propiedades de la función indicadora.

Sea  $\Omega$  cualquier espacio y  $\mathcal{A}$  cualquier colección de subconjuntos de  $\Omega$ ,

- (i)  $I_A(\omega) = 1 - I_{A^c}(\omega)$  para todo  $A \in \mathcal{A}$
- (ii)  $I_{A_1 \cap A_2 \cap \dots \cap A_n}(\omega) = I_{A_1}(\omega) \cdot I_{A_2}(\omega) \cdots I_{A_n}(\omega)$  con  $A_1, A_2, \dots, A_n \in \mathcal{A}$
- (iii)  $I_{A_1 \cup A_2 \cup \dots \cup A_n}(\omega) = \max[I_{A_1}(\omega), I_{A_2}(\omega), \dots, I_{A_n}(\omega)]$  con  $A_1, A_2, \dots, A_n \in \mathcal{A}$
- (iv)  $I_A^2(\omega) = I_A(\omega)$  para todo  $A \in \mathcal{A}$

La demostración de las anteriores propiedades se dejan como ejercicio al lector.

La función indicadora por lo general será usada para denotar subconjuntos de la recta real, por ejemplo:

$$I_{(0,1)}(x) = I_{(0,1)}(x) = \begin{cases} 1 & \text{si } 0 \leq x < 1 \\ 0 & \text{en otro caso (e.o.c.)} \end{cases}$$

$$I_{\mathbb{Z}^+}(x) = \begin{cases} 1 & \text{si } x \in \text{enteros positivos} \\ 0 & \text{e.o.c.} \end{cases}$$

La función indicadora será frecuentemente usada en lo que resta del presente documento. A menudo se utiliza para simplificar la notación de funciones, como por ejemplo:

### Ejemplo 15.

Sea la función  $f(\cdot)$  descrita por

$$f(x) = \begin{cases} 0 & \text{si } 0 \leq x \\ x & \text{si } 0 < x \leq 1 \\ 2-x & \text{si } 1 < x \leq 2 \\ 0 & \text{si } 2 < x \end{cases}$$

utilizando la función indicadora se puede escribir como

$$f(x) = x \cdot I_{(0,1)}(x) + (2-x) \cdot I_{(1,2)}(x).$$

Otro tipo de funciones que serán empleadas en futuras discusiones son las llamadas funciones conjunto:

### Definición 14, Funciones de Conjunto.

Cualquier función cuyo dominio sea una colección de conjuntos y su contradominio la línea real incluyendo  $\pm\infty$ , es una *función de conjunto*.

Son de especial interés por que la función de probabilidad es un tipo particular de función de conjunto. Listamos ejemplos a continuación:

#### Ejemplo 16.

Sea  $\Omega$  el espacio muestral del experimento que consiste en lanzar dos dados, y sea  $\mathcal{A}$  la colección de todos los subconjuntos de  $\Omega$ . Definimos la función  $N(A) =$  número de puntos (resultados) del espacio muestral  $\Omega$  con  $A \in \mathcal{A}$ . Entonces  $N(\emptyset) = 0$ ,  $N(\Omega) = 36$ ,  $N(\{A \in \mathcal{A} \mid \text{la suma de los dados sea igual a } 7\}) = 6$ .

La función definida en el ejemplo anterior es conocida por función *tamaño del conjunto* y es válida para todo conjunto  $A$  de cualquier colección de conjuntos  $\mathcal{A}$ .

#### Ejemplo 17.

Sea  $\Omega$  el plano cartesiano y  $\mathcal{A}$  cualquier colección de subconjuntos de  $\Omega$  para los cuales se pueda asignar un área. Entonces definimos  $Q(A) =$  Área de  $A$ , con  $A \in \mathcal{A}$ . Por ejemplo, si  $A = \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}$ , entonces  $Q(A) = 1$ ; si  $A = \{(x, y) \mid x^2 + y^2 = r^2\}$  entonces  $Q(A) = 2\pi r^2$ ; si  $A = \{(0, 0), (1, 1)\}$ , entonces  $Q(A) = 0$ .

Otro concepto necesario para poder definir a la función de probabilidad es un tipo particular de álgebra booleana:

### Definición 15, Sigma Álgebra.

Una *sigma álgebra* es una álgebra o álgebra Booleana la cual además cumple con la siguiente propiedad:

Si  $A_1, A_2, \dots \in \mathcal{A}$ , entonces  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ .

Por lo tanto, toda sigma álgebra es a su vez una álgebra, pero no toda álgebra es una sigma álgebra.

Por otra parte, el espacio de eventos de un experimento es una sigma álgebra. Por consiguiente, cualquier propiedad o definición sobre una sigma álgebra implica que se cumple también para el espacio de eventos. Tal es el caso de la definición de función de probabilidad.

Ahora sí contamos con todos los elementos necesarios para poder dar la definición formal de probabilidad.

### Definición 16, Función de Probabilidad.

Sea  $\Omega$  el espacio muestral de un experimento aleatorio y sea  $\mathcal{A}$  una colección de eventos del experimento que además cumpla con ser una sigma álgebra. Entonces una función de probabilidad  $P[\cdot]$  es una función de conjunto con dominio en  $\mathcal{A}$  y contradominio en el intervalo  $[0, 1]$  que satisface los siguientes axiomas:

- (i)  $P[A] \geq 0$  para todo  $A \in \mathcal{A}$ .
- (ii)  $P[\Omega] = 1$ .
- (iii) Si  $A_1, A_2, \dots \in \mathcal{A}$  son mutuamente excluyentes, es decir,  $A_i \cap A_j = \emptyset$  para todo  $i \neq j$  con  $i, j = 1, 2, \dots$ , entonces se tiene que  $P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i]$ .

Ciertamente estos axiomas se inspiraron en las definiciones de probabilidad clásica y probabilidad frecuentista. La definición formal de probabilidad es una definición matemática; nos dice cual es el conjunto de funciones que pueden ser llamadas funciones de probabilidad, pero no nos dice que valor específico asigna  $P[\cdot]$  a un evento  $A$ . Tendremos que moldear o manipular nuestro experimento aleatorio de tal manera que nos permita obtener valores para la probabilidad de eventos.

$P[A]$  se lee como "la probabilidad del evento  $A$ " o "la probabilidad que el evento  $A$  ocurra" lo cual quiere decir la probabilidad de que cualquier resultado contenido en  $A$  ocurra. Entonces, calcular la probabilidad que ocurra cualquier resultado de un subconjunto de resultados del experimento, es equivalente a calcular la probabilidad del evento que los relaciona.

Para el caso de la función de probabilidad, usamos corchetes en vez de paréntesis para encerrar el elemento del dominio en que se está aplicando la función.

### Ejemplo 18.

Considere el experimento de lanzar dos monedas, digamos una de 50 centavos y una de un peso. Sea  $\Omega = \{(S, S), (S, A), (A, S), (A, A)\}$  donde el primer componente de la tupla  $(\cdot, \cdot)$  representa el resultado de la moneda de 50 centavos. Modelemos este experimento aleatorio suponiendo que los cuatro

elementos de  $\Omega$  son equiprobables, es decir  $P(\{(S, S)\}) = P(\{(S, A)\}) = P(\{(A, S)\}) = P(\{(A, A)\})$ . Surge entonces la siguiente pregunta: ¿La función  $P[\cdot]$  implícitamente definida en éste ejemplo una función de probabilidad, es decir, cumple con los 3 axiomas de la definición de probabilidad? Se puede mostrar que sí cumple con los 3 axiomas y que por lo tanto es una función de probabilidad.

En nuestra definición de evento y  $\mathcal{A}$  una colección de eventos, hemos afirmamos que  $\mathcal{A}$  no siempre puede ser tomada como la colección de todos los subconjuntos de  $\Omega$ . La razón para esto es que existen espacios muestrales  $\Omega$  para los cuales la colección de todos los subconjuntos de  $\Omega$  es tan grande que es imposible definir una función de probabilidad que cumpla con los tres axiomas de la definición.

Ahora somos capaces de deducir algunas otras propiedades de nuestra función  $P[\cdot]$  que se desprenden de la definición y de los tres axiomas. Es al enunciar y al probar éstas propiedades que se hará evidente la conveniencia de asumir a  $\mathcal{A}$  como una sigma álgebra de eventos. Puesto que  $\mathcal{A}$  es el dominio de  $P[\cdot]$ , solo elementos de  $\mathcal{A}$  pueden ser puestos en lugar del punto en la notación  $P[\cdot]$ . Por ser  $\mathcal{A}$  un álgebra, asumimos que si  $A, B \in \mathcal{A}$ , entonces  $A^c, A \cup B, A \cap B, A^c \cup B$ , etc.  $\in \mathcal{A}$ , y por lo tanto será válido también hablar de las probabilidades  $P[A^c], P[A \cup B], P[A \cap B], P[A^c \cup B]$ , etc.

### Propiedades de $P[\cdot]$ .

En cada uno de los siguientes teoremas, se asume que  $\Omega$  y  $\mathcal{A}$  son dados y  $P[\cdot]$  es una función de probabilidad con dominio  $\mathcal{A}$ .

- **Teorema 15,  $P[\emptyset] = 0$ .**

#### Demostración.

Escoja  $A_1 = \emptyset, A_2 = \emptyset, A_3 = \emptyset, \dots$ ; por el axioma (I) se tiene que

$$P[\emptyset] = P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i] = \sum_{i=1}^{\infty} P[\emptyset], \text{ es decir,}$$

$$P[\emptyset] = \sum_{i=1}^{\infty} P[\emptyset], \text{ lo cual se cumple si y solo si}$$

$$P[\emptyset] = 0$$

• **Teorema 16.**

Si  $A_1, A_2, \dots, A_n$  son mutuamente excluyentes en  $\mathcal{A}$  entonces

$$P[A_1 \cup \dots \cup A_n] = \sum_{i=1}^n P[A_i]$$

**Demostración.**

Escoja  $A_{n+1} = \emptyset, A_{n+2} = \emptyset, \dots$ ; entonces

$$\begin{aligned} \bigcup_{i=1}^n A_i &= \left( \bigcup_{i=1}^n A_i \right) \cup \left( \bigcup_{i=n+1}^{\infty} A_i \right) = \left( \bigcup_{i=1}^n A_i \right) \cup \left( \bigcup_{i=n+1}^{\infty} \emptyset \right) = \left( \bigcup_{i=1}^n A_i \right) \cup \emptyset = \left( \bigcup_{i=1}^n A_i \right) \\ \therefore P\left[ \bigcup_{i=1}^n A_i \right] &= P\left[ \bigcup_{i=1}^{\infty} A_i \right] = \sum_{i=1}^{\infty} P[A_i] = \sum_{i=1}^n P[A_i] + \sum_{i=n+1}^{\infty} \emptyset = \sum_{i=1}^n P[A_i] + \sum_{i=n+1}^{\infty} 0 = \\ &\Rightarrow P\left[ \bigcup_{i=1}^n A_i \right] = \sum_{i=1}^n P[A_i] \quad \blacksquare \end{aligned}$$

• **Teorema 17.**

Si  $A \in \mathcal{A}$ , entonces,  $P[A^c] = 1 - P[A]$

**Demostración.**

$$\begin{aligned} &\text{Puesto que } A \cup A^c = \Omega, \text{ y } A \cap A^c = \emptyset, \\ \Rightarrow P[\Omega] &= P[A \cup A^c] = P[A] + P[A^c] \\ &\text{y por el axioma (ii) tenemos que } P[\Omega] = 1, \\ \Rightarrow 1 &= P[A] + P[A^c] \\ \Rightarrow P[A^c] &= 1 - P[A] \quad \blacksquare \end{aligned}$$

• **Teorema 18.**

Si  $A, B \in \mathcal{A}$ , entonces,

$$\begin{aligned} P[A] &= P[A \cap B] + P[A \cap B^c] \text{ y} \\ P[A - B] &= P[A \cap B^c] = P[A] - P[A \cap B]. \end{aligned}$$

**Demostración.**

$$\begin{aligned} &\text{Ya que } A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c), \\ &\text{y } (A \cap B) \cap (A \cap B^c) = (A \cap A) \cap (B \cap B^c) = A \cap \emptyset = \emptyset \text{ (excluyentes)} \\ \Rightarrow P[A] &= P[(A \cap B) \cup (A \cap B^c)] = P[(A \cap B)] + P[(A \cap B^c)] \\ &\text{si despejamos} \quad \blacktriangleright \end{aligned}$$

$$P[A - B] = P[A \cap B^c] = P[A] - P[(A \cap B)]$$

● **Teorema 19.**

Si  $A, B \in \mathcal{A}$ , entonces  $P[A \cup B] = P[A] + P[B] - P[A \cap B]$ .

Más general aún, si  $A_1, A_2, \dots, A_n \in \mathcal{A}$ , entonces

$$P[A_1 \cup A_2 \cup \dots \cup A_n] =$$

$$\sum_{j=1}^n P[A_j] - \sum_{1 \leq i < j} P[A_i \cap A_j] + \sum_{1 \leq i < j < k} P[A_i \cap A_j \cap A_k] - \dots \\ \dots + (-1)^{n+1} P[A_1 \cap A_2 \cap \dots \cap A_n]$$

**Demostración.**

$$A \cup (A^c \cap B) = (A \cup A^c) \cap (A \cup B) = \Omega \cap (A \cup B) = A \cup B, \text{ y}$$

$$A \cap (A^c \cap B) = (A \cap A^c) \cap B = \emptyset \cap B = \emptyset \text{ (excluyentes),}$$

$$\Rightarrow P[A \cup B] = P[A \cup (A^c \cap B)] = P[A] + P[(A^c \cap B)], \text{ y por teorema 18}$$

$$\Rightarrow P[A \cup B] = P[A] + (P[B] - P[B \cap A])$$

La parte más general del teorema se demuestra por inducción matemática. ▸

● **Teorema 20.**

Si  $A, B \in \mathcal{A}$ , y  $A \subset B$ , entonces  $P[A] \leq P[B]$

**Demostración.**

$$B = (B \cap A) \cup (B \cap A^c) \text{ y } B \cap A = A$$

$$\Rightarrow B = (A) \cup (B \cap A^c), \text{ y como son excluyentes pues } (A) \cap (B \cap A^c) = \emptyset$$

$$\Rightarrow P[B] = P[A] + P[B \cap A^c], \text{ y como } P[A] \geq 0$$

$$\Rightarrow P[B] \geq P[B \cap A^c]$$

● **Teorema 21, Desigualdad de Boole**

Si  $A_1, A_2, \dots, A_n \in \mathcal{A}$ , entonces

$$P[A_1 \cup A_2 \cup \dots \cup A_n] \leq P[A_1] + P[A_2] + \dots + P[A_n]$$

**Demostración.**

Para  $A_1$  y  $A_2$  se tiene que

$$P[A_1 \cup A_2] = P[A_1] + P[A_2] - P[A_1 \cap A_2] \leq P[A_1] + P[A_2]$$

La demostración se completa con inducción matemática

### **Definición 17, Espacio de Probabilidad.**

Un *espacio de probabilidad* es la triplete  $(\Omega, \mathcal{A}; P[\cdot])$  donde  $\Omega$  es el espacio muestral,  $\mathcal{A}$  es una sigma álgebra de eventos de  $\Omega$ , y  $P[\cdot]$  es una función de probabilidad con dominio  $\mathcal{A}$ .

El único objetivo de definir el espacio de probabilidad, es contar con un término que de manera clara y correcta implique la existencia de los tres componentes de su notación. Así en el futuro contaremos con un término que simplificará nuestros teoremas, definiciones, etc.

### **3.4 Espacios Muestrales Finitos.**

En la sección anterior definimos formalmente espacio muestral, evento y probabilidad, culminando con la definición de espacio de probabilidad. Aclaramos también que tales definiciones no eran suficientes por sí solas, para alcanzar nuestro objetivo que es calcular la probabilidad que tiene un evento de ocurrir; que para lograr tal objetivo además tenemos que moldear adecuadamente el experimento. En esta sección mostraremos como puede hacerse esto para espacios muestrales finitos, esto es, espacios muestrales con un número finito de puntos o elementos.

En cierto tipo de problemas, de los cuales los juegos de azar son notables ejemplos, el espacio muestral contiene un número finito de puntos, digamos  $N = N(\Omega)$  ( $N(\cdot)$  es la función de conjunto que denota el número de elementos de un conjunto). Algunos de éstos problemas pueden ser modelados asumiendo que los puntos del espacio muestral son equiprobables. Tales problemas son el tema de la siguiente discusión.

#### **Espacios muestrales finitos con puntos equiprobables.**

Para algunos experimentos aleatorios hay un número finito de resultados, digamos  $N$ , y con frecuencia es aceptable asumir que la probabilidad de cada resultado es  $1/N$ . La definición de probabilidad clásica es por lo general adecuada en éste tipo de problemas, pero veremos como la definición axiomática es aplicable también. Sean  $\omega_1, \omega_2, \dots, \omega_N$  los  $N$  puntos del espacio muestral  $\Omega$ . Sea  $P[\cdot]$  la función de conjunto que tiene como dominio a la colección de todos los subconjuntos de  $\Omega$  y que satisface las siguientes condiciones:

$$(i) \quad P\{\{\omega_1\}\} = P\{\{\omega_2\}\} = \dots = P\{\{\omega_N\}\}.$$

- (ii) Si  $A$  es cualquier subconjunto de  $\Omega$  que contiene  $N(A)$  puntos, entonces  $P[A] = N(A)/N$ .

La función de conjunto  $P[\cdot]$  así definida cumple con los tres axiomas de la definición de probabilidad y por lo tanto es una función de probabilidad.

Nótese que la función  $P[\cdot]$  antes definida, recrea la definición de probabilidad clásica bajo la formal estructura de la definición axiomática de probabilidad.

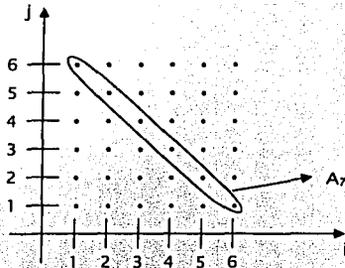
### **Definición 18, Función de Probabilidad Equiprobable.**

La función de probabilidad  $P[\cdot]$  que satisface las anteriores condiciones (i) y (ii) es llamada *función de probabilidad equiprobable*.

Una vez determinado que un experimento aleatorio puede ser manipulado bajo el supuesto que sus puntos muestrales son equiprobables, entonces el único problema para determinar la probabilidad de que el evento  $A$  ocurra es determinar  $N(A)$  y  $N(\Omega)$ . El problema se reduce a contar el número de elementos de  $A$  y el número de elementos en  $\Omega$ .

### **Ejemplo 19.**

Considere el experimento de lanzar dos dados. Entonces  $\Omega = \{(i, j) \mid i, j = 1, 2, \dots, 6\}$ , donde  $i$  denota el resultado del dado 1 y  $j$  el resultado del dado 2. En total hay 6-6 puntos muestrales, es decir,  $N(\Omega) = 36$ . Puesto que los dados están aceptablemente balanceados, es razonable asignar  $1/36$  a la probabilidad de cada punto muestral.  $\Omega$  puede ser representado por medio de una cuadrícula como lo muestra la siguiente figura. Sea el evento  $A_7 = \{\text{que la suma de los dados sea } 7 = \{(i, j) \mid i, j = 1, 2, \dots, 6 \text{ y } i + j = 7\}\}$ .  $A_7$  es explícitamente descrito por  $\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ ; entonces  $N(A_7) = 6$ . Por lo tanto  $P[A_7] = N(A_7)/N(\Omega) = 6/36 = 1/6$ . Similarmente para los eventos  $A_j = \{\text{los dados suma } j\}$ , con  $j = 2, 3, \dots, 12$ ,  $P[A_j]$  puede ser fácilmente calculada.



Si  $N(A)$  y  $N(\Omega)$  son grandes para un experimento aleatorio dado, con un número finito de resultados equiprobables, el conteo de los mismos se puede convertir en un problema. Tal conteo puede facilitarse a menudo usando fórmulas de cálculo combinatorio, algunas de las cuales se desarrollarán a continuación.

### Muestreo.

Piense en un experimento en el que, debido a su naturaleza, cada resultado puede ser representado mediante una tupla o  $n$ -ada. Este es el caso del ejemplo anterior, en el que cada resultado es representado por una tupla de 2 entradas. Frecuentemente se utiliza la terminología de *muestreo* para describir un simple experimento aleatorio de urnas y extracción de pelotas. Suponga una urna que contiene  $M$  pelotas numeradas de la 1 a la  $M$ . El experimento consiste en sacar, una por una,  $n$  pelotas de la urna; se dice que se tiene una muestra de tamaño  $n$ . La extracción debe realizarse de tal manera, que cualquier pelota en la urna tenga la misma posibilidad de ser seleccionada que las demás; decimos que la pelota ha sido seleccionada aleatoriamente.

Son *con reemplazo* y *sin reemplazo* los dos principales tipos de muestreo. Se dice que una muestra se obtiene con reemplazo, si después de extraer cada pelota esta es registrada y devuelta a la urna. Por el contrario, se dice que una muestra se obtiene sin reemplazo, si después de extraer cada pelota esta no se devuelve a la urna. Por lo tanto, en el muestreo sin reemplazo se tiene que  $n$  debe ser menor o igual a  $M$ , mientras que en el muestreo con reemplazo  $n$  puede ser cualquier número entero positivo incluso mayor a  $M$ . Los resultados de la extracción de muestras de tamaño  $n$  se pueden registrar por medio de tuplas de

la forma  $(z_1, z_2, \dots, z_n)$ , donde  $z_i$  denota el número de la bola en la  $i$ -ésima extracción.

Si recordamos la función de probabilidad definida para este tipo de experimentos, debemos calcular el tamaño ( $N(A)$ ) del evento  $A$  representado mediante tuplas con ciertas características en común. Tal número se puede obtener mediante el siguiente procedimiento: Primero determine el número de objetos, digamos  $N_1$ , que pueden ser usados en la primera posición de la tupla. Después determine el número de objetos, digamos  $N_2$ , que pueden ser usados en la segunda posición de la tupla, asumiendo que el resultado de la primera entrada es conocido. De manera similar, determine el número de objetos, digamos  $N_3$ , que pueden ser usados en la tercera posición de la tupla, asumiendo que los resultados de la primera y segunda entradas son conocidos. Continúe así hasta que  $N_n$  halla sido determinado. Entonces el tamaño  $N(A)$  o número de tuplas del evento  $A$  es igual a  $N_1 \cdot N_2 \cdot \dots \cdot N_n$ .

### Ejemplo 20, Ordenaciones.

El número total de diferentes muestras ordenadas con  $n$  bolas que se pueden obtener de una urna que contiene  $M$  bolas diferentes, es igual a  $M^n$  si la extracción es realizada con reemplazo y  $M(M-1) \dots (M-(n-1))$  si la extracción es realizada sin reemplazo. Una muestra ordenada puede ser representada por  $(z_1, z_2, \dots, z_n)$ , donde  $z_i$  denota el número de la bola en la  $i$ -ésima extracción.

En el muestreo con reemplazo, hay  $M$  bolas que pueden ser seleccionadas en la primera entrada de la tupla,  $M$  bolas en la segunda entrada, así igual para cualquier otra entrada de la tupla. Por lo tanto hay un total de  $M^n$  diferentes tuplas.

En el muestreo sin reemplazo, hay  $M$  bolas que pueden ser seleccionadas en la primera entrada de la tupla,  $M-1$  bolas en la segunda entrada, así sucesivamente hasta  $M-(n-1)$  bolas en la entrada  $n$  de la tupla. Por lo tanto hay un total de  $M(M-1) \dots (M-n+1)$  diferentes tuplas. Para abreviar, el número  $M(M-1) \dots (M-n+1)$  se puede denotar como  $(M)_n$  ó  $O_{(M,n)}$  ó  $O_M^n$  y se lee como "las ordenaciones de  $M$  elementos tomados de  $n$  en  $n$ ".

$$O_M^n = M \cdot (M-1) \cdot \dots \cdot (M-n+1) = \frac{M!}{(M-n)!} \quad (AN1)$$

## Ejemplo 21, Combinaciones.

Sea  $S$  cualquier conjunto que contiene  $M$  elementos, ¿cuántos subconjuntos diferentes de tamaño  $n$  se pueden formar con los elementos de  $S$ ? El número que contesta tal pregunta es conocido como "las *combinaciones* de  $M$  elementos tomados de  $n$  en  $n$ " y se denota por  $\binom{M}{n}$  o  $x_n$ . Ya se ha explicado que un

subconjunto de tamaño  $n$ , es una colección de  $n$  objetos los cuales no están ordenados. Por ejemplo, el subconjunto  $\{S_1, S_5, S_7\}$  es el mismo subconjunto que  $\{S_5, S_1, S_7\}$  puesto que contienen los mismos 3 elementos.

No obstante, en este problema no importa el orden de los elementos, haremos uso del conocimiento que tenemos sobre tuplas ordenadas para deducir cuánto vale  $x_n$ . Sin conocer aún  $x_n$ , sabemos que cada subconjunto de tamaño  $n$  se puede ordenar de  $O_{(n, n)} = n!$  maneras distintas, pues se trata de muestreo sin reemplazo. Por lo tanto hay en total  $n! \cdot x_n$  diferentes muestras ordenadas que se pueden formar con los  $x_n$  diferentes subconjuntos, número que debe ser igual a  $O_{(M, n)}$ . Se obtiene así la siguiente igualdad  $n! \cdot x_n = O_{(M, n)}$  de la cual podemos despejar  $x_n$ :

$$x_n = \binom{M}{n} = \frac{O_{(M, n)}}{n!} = \frac{M!}{(M-n)!n!} \quad (\text{AN2})$$

$$\text{Teorema 22, } \binom{M}{n} = \binom{M}{M-n}.$$

### Demostración

$$\binom{M}{M-n} = \frac{M!}{(M-(M-n))!(M-n)!} = \frac{M!}{(n)!(M-n)!} = \frac{M!}{(M-n)!n!} = \binom{M}{n}$$

## Ejemplo 22, Tamaño de Espacio de Eventos.

Si recordamos, previamente afirmamos que el número de elementos del espacio de eventos para un  $\Omega$  de tamaño  $n$ , está dado por  $2^n$ . Ahora ya podemos demostrar tal afirmación usando el concepto de combinaciones:

$$\binom{M}{n} = \text{Número de subconjuntos diferentes de tamaño } n \text{ que se pueden formar con } M \text{ elementos.}$$

$$\Rightarrow \sum_{n=0}^M \binom{M}{n} = \text{Es el total de diferentes subconjuntos que se pueden formar con } m \text{ elementos} \\ = N(\text{Espacio de Eventos})$$

$$\text{Por el Binomio de Newton sabemos que } (a+b)^M = \sum_{n=0}^M \binom{M}{n} a^n b^{M-n}$$

$$\Rightarrow \sum_{n=0}^M \binom{M}{n} = \sum_{n=0}^M \binom{M}{n} \cdot 1^n \cdot 1^{M-n} = (1+1)^M = 2^M. \quad \blacksquare \quad (\text{AN3})$$

Es común cuando se resuelven problemas con cálculo combinatorio, confundirse y utilizar combinaciones en vez de ordenaciones o viceversa. Recuerde el lector que las ordenaciones hacen referencia al número de tuplas ordenadas o elementos de  $\Omega$ , mientras que las combinaciones hacen referencia al número de subconjuntos no ordenados de  $\Omega$ . El siguiente ejemplo es una prueba de fuego al manejo de ambos conceptos.

### Ejemplo 23.

Una urna tiene  $M$  bolas numeradas de 1 a  $M$ , de las cuales las primeras  $K$  son defectuosas y las restantes  $M - K$  no son defectuosas. El experimento consiste en sacar  $n$  bolas de la urna. Sea el evento  $A_r = \{\text{exactamente salen } r \text{ bolas defectuosas}\}$ . ¿Cuánto vale  $P(A_r)$  cuando las bolas se extraen (i) con reemplazo y (ii) sin reemplazo?

Sea  $\Omega = \{(z_1, \dots, z_n) \mid z_j \text{ es el número de la bola obtenida de la } j\text{-ésima extracción}\}$ .

$$\text{Sabemos que } P(A_r) = \frac{N(A_r)}{N(\Omega)}$$

Por el ejemplo 20, tenemos que bajo (i)  $N(\Omega) = O_{(M, n)}$  y bajo (ii)  $N(\Omega) = M^n$ .  $A_r$  es el subconjunto de los elementos de  $\Omega$  para los cuales  $z_j$  es un número entre 1 y  $K$ , es decir,  $A_r = \{(z_1, \dots, z_n) \in \Omega \mid 1 \leq z_j \leq K \forall 1 \leq j \leq n\}$ . Estas  $r$  bolas deben caer en algún subconjunto de  $r$  posiciones de entre un total de  $n$  posiciones posibles. Hay entonces  $\binom{n}{r}$  maneras diferentes de seleccionar las  $r$  posiciones (no bolas) para las bolas con número entre 1 y  $K$ .

Ahora, para cada una de las  $\binom{n}{r}$  diferentes posiciones, existen  $K^r (M - K)^{n-r}$  diferentes tuplas ordenadas para (i) y  $O_{(k, k)} \cdot O_{(M-k, n-k)}$  diferentes tuplas

ordenadas para (II). Entonces para (I)  $N(A_r) = \binom{n}{r} k^r (M-K)^{n-r}$  y para (II)  $N(A_r) = \binom{n}{r} O_{(K,r)} O_{(M-K,n-r)}$ . Así se tiene que

$$P_{\{A_r\}} = \frac{\binom{n}{r} k^r (M-K)^{n-r}}{M^n} \text{ para (I) y} \quad (AN4)$$

$$P_{\{A_r\}} = \frac{\binom{n}{r} O_{(K,r)} O_{(M-K,n-r)}}{O_{(M,n)}} \quad (AN5)$$

El resultado anterior (AN5) se puede escribir también de la siguiente manera,

$$P_{\{A_r\}} = \frac{\binom{K}{r} \binom{M-K}{n-r}}{\binom{M}{n}}. \quad (AN6)$$

Por lo ilustrativo que resulta, a continuación se deducirá (AN6) utilizando subconjuntos y no tupla ordenadas, es decir, combinaciones en vez de ordenaciones.

Suponga ahora, que nuestro espacio muestral  $\Omega'$  está formado por subconjuntos de tamaño  $n$ , en vez de tuplas ordenadas.  $\Omega' = \{\{z_1, z_2, \dots, z_n\} \mid z_1, z_2, \dots, z_n \text{ son los números de las } n \text{ bolas extraídas}\}$ . Uno podría pensar en  $\Omega'$  como consecuencia de extraer las  $n$  bolas de un solo en vez de una por una.

Por la definición de combinaciones, sabemos que hay  $\binom{M}{n}$  subconjuntos de tamaño  $n$  que se pueden formar con  $M$  bolas, por lo que  $N(\Omega') = \binom{M}{n}$ . Si se asume que estos subconjuntos son equiprobables, se puede decir que  $P\{A'_r\} = \frac{N(A'_r)}{N(\Omega')}$ .

Ahora  $N(A'_r)$  es el tamaño del evento que consta de todos aquellos subconjuntos de tamaño  $n$  los cuales tienen exactamente  $r$  bolas de las numeradas entre 1 y  $K$ .

Las  $r$  bolas de entre 1 a  $K$  se pueden seleccionar de  $\binom{K}{r}$  diferentes maneras y las restantes  $n - r$  bolas de  $\binom{M-K}{n-r}$ . Por lo tanto  $N(A'_r) = \binom{K}{r} \binom{M-k}{n-r}$  y finalmente

$$P[A'_r] = \frac{N(A'_r)}{N(\Omega)} = \frac{\binom{K}{r} \binom{M-K}{n-r}}{\binom{M}{n}}$$

Se demostró que la probabilidad de seleccionar exactamente  $r$  elementos defectuosos sin reemplazo, es igual considerando dos diferentes espacios muestrales; uno que consiste de tuplas ordenadas y otro que consiste de subconjuntos de tamaño  $n$ .

Una sencilla manera para recordar la fórmula (AN6) es notar que  $(K) + (M - K) = M$  y  $(r) + (n - r) = n$ , es decir, la suma de los términos superiores del numerador es igual al término superior del denominador y la suma de los términos inferiores del numerador es igual al término inferior del denominador.

### Ejemplo 24.

La ecuación (AN6) es particularmente útil para calcular algunas probabilidades que involucran juegos con cartas. Por ejemplo, podríamos preguntarnos por la probabilidad que en una "mano" de 13 cartas contenga exactamente 6 espadas? En total hay 52 cartas y 13 cartas espadas (bolas defectuosas haciendo una analogía con la terminología de muestreo) en una baraja, y se puede diseñar el proceso de revolver y repartir las cartas asumiendo que cada mano es una muestra de tamaño 13 sin reemplazo. Sea el evento  $A_6 =$  {salen exactamente 6 cartas espadas}, por (AN6) tenemos que

$$P[A_6] = \frac{N(A_6)}{N(\Omega)} = \frac{\binom{13}{6} \binom{52-13}{13-6}}{\binom{52}{13}}$$

Existen otras fórmulas para calcular probabilidades en experimentos con espacios muestrales finitos de elementos equiprobables, las cuales pueden ser obtenidas usando los métodos del análisis combinatorio. Sin embargo, ya no nos encargaremos de ello.

### Espacios muestrales finitos con puntos no equiprobables.

Vimos que en los espacios muestrales finitos con elementos equiprobables, la función de probabilidad de cualquier evento  $A$  esta dada por  $P[A] = \frac{N(A)}{N(\Omega)}$ . Para

los espacios muestrales finitos con elementos no equiprobables, las cosas no son tan sencillas. Sin embargo, siempre podremos definir el valor  $P[A]$  para cada uno de los  $2^{N(\Omega)}$  eventos del espacio de eventos, especificando el valor de  $P[\cdot]$  para cada uno de los  $N(\Omega)=N$  eventos elementales. Sea  $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$  y  $p_j = P[\{\omega_j\}]$  con  $1 \leq j \leq N$ . Puesto que

$$1 = P[\Omega] = P\left[\bigcup_{j=1}^N \{\omega_j\}\right] = \sum_{j=1}^N P[\{\omega_j\}] = \sum_{j=1}^N p_j,$$

se define  $P[\cdot]$  para todo evento  $A$  como

$$P[A] = P\left[\bigcup_{\omega_j \in A} \{\omega_j\}\right] = \sum p_j \quad \omega_j \in A.$$

Se puede demostrar que la función de probabilidad arriba definida satisface los tres axiomas de la definición de función de probabilidad.

### Ejemplo 25.

Considere un experimento que tiene  $N$  posibles resultados, digamos  $\omega_1, \omega_2, \dots, \omega_N$ , donde cada resultado  $\omega_{j+1}$  es dos veces más probable que el resultado  $\omega_j$  con  $j=1, \dots, N-1$ , es decir,  $p_{j+1} = 2p_j$  donde  $p_i = P[\{\omega_i\}]$ . Determine  $P[A_k]$  donde  $A_k = \{\omega_1, \omega_2, \dots, \omega_k\}$ . Ya que

$$\begin{aligned} \sum_{j=1}^N p_j &= \sum_{j=1}^N 2^{j-1} p_1 = p_1 \sum_{j=1}^N 2^{j-1} = p_1 (1 + 2 + 2^2 + \dots + 2^{N-1}) = p_1 (2^N - 1) = 1 \\ \Rightarrow p_1 &= \frac{1}{2^N - 1} \\ \therefore p_j &= \frac{2^{j-1}}{2^N - 1} \end{aligned}$$

$$\text{Entonces } P[A_k] = \sum_{j=1}^k p_j = \sum_{j=1}^k \frac{2^{j-1}}{2^N - 1} = \frac{1}{2^N - 1} \sum_{j=1}^k 2^{j-1} = \frac{2^k - 1}{2^N - 1}.$$

### 3.5 Probabilidad Condicional e Independencia.

Al aplicar la teoría de probabilidad en problemas prácticos, es frecuente que el experimentador se encuentre con situaciones del siguiente tipo: dado que algo ocurre, cual es la probabilidad de que algo más ocurra? Por ejemplo, en el experimento de registrar el tiempo de vida de un foco, uno pudiera preguntarse por la probabilidad de que un foco dure al menos 100 horas, dado que ya ha durado 24 horas. En un experimento de sacar muestras de una caja que contiene 100 resistencias de las cuales 5 son defectuosas, cual es la probabilidad de que

en la tercera extracción se obtenga una resistencia defectuosa dado que las dos primeras resultaron defectuosas? Preguntas de probabilidad de este tipo son consideradas del campo de estudio de la probabilidad condicional, tema que a continuación estudiaremos.

### Probabilidad Condicional.

Empezamos asumiendo que trabajamos en un espacio probabilístico  $(\Omega, \mathcal{A}, P[\cdot])$ , es decir, tenemos algún experimento aleatorio en el cual el espacio muestral  $\Omega$ , la colección de eventos  $\mathcal{A}$  y la función de probabilidad  $P[\cdot]$  han sido definidas.

Dados dos eventos A y B, queremos definir la probabilidad que tiene A de ocurrir dado que B ha ocurrido.

### Definición 19, Probabilidad Condicional.

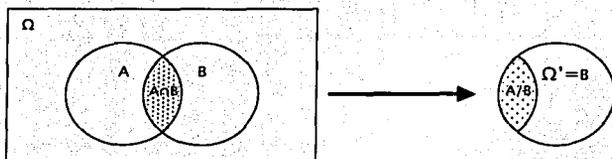
Sean A y B dos eventos en  $\mathcal{A}$  del espacio probabilístico  $(\Omega, \mathcal{A}, P[\cdot])$ . La *probabilidad condicional* del evento A dado el evento B, denotada por  $P[A/B]$ , esta dada por

$$P[A/B] = \frac{P[A \cap B]}{P[B]} \text{ con } P[B] > 0. \quad (\text{AN7})$$

### Nota.

Una fórmula que es evidente a partir de la anterior definición, es  $P[A \cap B] = P[A/B] \cdot P[B] = P[B/A] \cdot P[A]$  si  $P[A]$  y  $P[B] \neq 0$ . Esta fórmula relaciona  $P[A/B]$  con  $P[B/A]$  en términos de las probabilidades no condicionales  $P[A]$  y  $P[B]$ .

De la definición de probabilidad condicional, pudiera no haber mucha diferencia entre la intersección y la probabilidad condicional entre eventos. Se debe tener siempre presente, que mientras la intersección de eventos es relativa al espacio muestral, la probabilidad condicional es relativa exclusivamente al evento que esta condicionando. Es decir, la probabilidad condicional se debe entender como una reducción del espacio muestral. El siguiente diagrama puede ayudarnos a entender esta importante idea:



Quisiéramos notar que la definición de probabilidad condicional, es compatible con la aproximación frecuentista de probabilidad, pues si uno observa un gran número de ocurrencias, digamos  $N$ , de un experimento aleatorio en el cual los eventos  $A$  y  $B$  son definidos, entonces  $P[A/B]$  representa la proporción de ocurrencias en las cuales  $B$  ocurrió y  $A$  también ocurre. Es decir

$$P[A/B] = \frac{N_{A \cap B}}{N_B},$$

donde  $N_B$  denota el número de ocurrencias del eventos  $B$  y  $N_{A \cap B}$  denota el número de ocurrencias del evento  $A \cap B$ . Puesto que  $P[A \cap B] = N_{A \cap B}/N$  y  $P[B] = N_B/N$ , entonces

$$\frac{P[A \cap B]}{P[B]} = \frac{N_{A \cap B}/N}{N_B/N} = \frac{N_{A \cap B}}{N_B} = P[A/B] \text{ lo cual es consistente con nuestra}$$

definición.

### Ejemplo 26.

Sea  $\Omega$  un espacio muestral finito, sea la colección de todos los subconjuntos de  $\Omega$  que en este caso es igual al espacio de eventos, y  $P[\cdot]$  la función de probabilidad equiprobable. Si  $A, B \in \mathcal{A}$ , entonces

$$P[A/B] = \frac{P[A \cap B]}{P[B]} = \frac{N(A \cap B)/N(\Omega)}{N(B)/N(\Omega)} = \frac{N(A \cap B)}{N(B)},$$

donde  $N(B)$  es el tamaño de  $B$  y  $N(B) > 0$ .

### Ejemplo 27.

Considere el experimento de lanzar dos monedas.  $\Omega = \{(S, S), (S, A), (A, S), (A, A)\}$ . Suponga además que los puntos de  $\Omega$  son equiprobables. Encuentre (i) la probabilidad las dos modas caigan sol dado que la primera cayó sol y (ii) la probabilidad las dos modas caigan sol dado que al menos una cayó sol. Sean  $A_1 = \{\text{la primera moneda cae sol}\}$  y  $A_2 = \{\text{la segunda moneda cae sol}\}$  entonces

$$(i) P[A_1 \cap A_2 / A_1] = \frac{P[(A_1 \cap A_2) \cap A_1]}{P[A_1]} = \frac{P[A_1 \cap A_2]}{P[A_1]} = \frac{1/4}{1/2} = 1/2$$

$$(ii) P[A_1 \cap A_2 / A_1 \cup A_2] = \frac{P[(A_1 \cap A_2) \cap (A_1 \cup A_2)]}{P[A_1 \cup A_2]} = \frac{P[A_1 \cap A_2]}{P[A_1 \cup A_2]} = \frac{1/4}{3/4} = 1/3$$

Nótese que (i) y (ii) pudieron ser contestados con números específicos, pero para ello, fue necesario manipular el experimento y suponer que los puntos del espacio muestral eran equiprobables.

Como ya se dijo, la probabilidad condicional es una reducción del espacio muestral  $\Omega$  en un nuevo espacio muestral  $\Omega' = B$ . Cabe entonces hacerse la siguiente pregunta: Dado un evento  $B$  tal que  $P[B] > 0$ , ¿es  $P[\cdot/B]$  una función de probabilidad? En otras palabras, ¿ $P[\cdot/B]$  satisface los tres axiomas de la definición de probabilidad? Observe que si  $P[B] > 0$  entonces

$$(i) P[A/B] = P[A \cap B] / P[B] \geq 0 \quad \forall A \in \mathcal{A}$$

$$(ii) P[\Omega/B] = P[B \cap \Omega] / P[B] = P[B] / P[B] = 1,$$

(iii) Sean  $A_1, A_2, \dots \in \mathcal{A}$  mutuamente excluyentes y sea  $\mathcal{A}$  una sigma álgebra

$$\Rightarrow P\left[\bigcup_{i=1}^{\infty} A_i / B\right] = \frac{P\left[\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right]}{P[B]} = \frac{P\left[\bigcup_{i=1}^{\infty} (A_i \cap B)\right]}{P[B]} = \frac{\sum_{i=1}^{\infty} P[A_i \cap B]}{P[B]} = \sum_{i=1}^{\infty} P[A_i/B]$$

Por lo tanto,  $P[\cdot/B]$  con  $P[B] > 0$  es una función de probabilidad; lo cual justifica el término de "probabilidad" condicional para su definición:  $P[\cdot/B]$  cumple también con las propiedades de cualquier función de probabilidad.

### Propiedades de $P[\cdot/B]$ .

En cada uno de los siguientes teoremas, se asume el espacio probabilístico  $(\Omega, \mathcal{A}, P[\cdot])$  esta dado y que  $B \in \mathcal{A}$  cumple que  $P[B] > 0$ .

● **Teorema 23,  $P[\emptyset/B] = 0$ .**

● **Teorema 24.**

Si  $A_1, A_2, \dots, A_n$  son mutuamente excluyentes en  $\mathcal{A}$  entonces

$$P[A_1 \cup \dots \cup A_n / B] = \sum_{i=1}^n P[A_i / B].$$

- **Teorema 25.**

Si  $A \in \mathcal{A}$ , entonces,  $P[A^c / B] = 1 - P[A / B]$ .

- **Teorema 26.**

Si  $A_1, A_2 \in \mathcal{A}$  entonces,

$$P[A_1 / B] = P[A_1 \cap A_2 / B] + P[A_1 \cap A_2^c / B].$$

- **Teorema 27.**

Si  $A_1, A_2 \in \mathcal{A}$  entonces  $P[A_1 \cup A_2 / B] = P[A_1 / B] + P[A_2 / B] - P[A_1 \cap A_2 / B]$ .

- **Teorema 28.**

Si  $A_1, A_2 \in \mathcal{A}$  y  $A_1 \subset A_2$ , entonces  $P[A_1 / B] \leq P[A_2 / B]$

- **Teorema 29.**

Si  $A_1, A_2, \dots, A_n \in \mathcal{A}$  entonces

$$P[A_1 \cup A_2 \cup \dots \cup A_n / B] \leq P[A_1 / B] + P[A_2 / B] + \dots + P[A_n / B]$$

La demostración de los anteriores teoremas no es necesaria por ser  $P[\cdot / B]$  una función de probabilidad. Existen algunas otras fórmulas que son de gran utilidad que involucran probabilidades condicionales, por lo que serán enunciadas como teoremas.

### **Teorema 30, Teorema de la Probabilidad Total.**

Para un espacio probabilístico  $(\Omega, \mathcal{A}, P[\cdot])$  cualquiera, si  $B_1, B_2, \dots, B_n$  es una partición del espacio muestral, entonces se tiene que

$$P[A] = \sum_{j=1}^n P[A / B_j] \cdot P[B_j] \text{ para cualquier } A \in \mathcal{A}$$

Una *partición del espacio muestral* es una colección de eventos  $B_1, B_2, \dots, B_n \in \mathcal{A}$ , que satisfacen  $\bigcup_{j=1}^n B_j = \Omega$ ,  $B_i \cap B_j = \emptyset \forall i \neq j$  y  $P[B_j] > 0 \forall 1 \leq j \leq n$ .

### Demostración

Ya que  $A = \bigcup_{j=1}^n A \cap B_j$  y que los  $(A \cap B_j)$  son mutuamente excluyentes se tenemos que

$$P[A] = P\left[\bigcup_{j=1}^n A \cap B_j\right] = \sum_{j=1}^n P[A \cap B_j] = \sum_{j=1}^n P[A/B_j]P[B_j]. \quad \blacksquare$$

### Corolario.

Para un espacio probabilístico dado  $(\Omega, \mathcal{A}, P[\cdot])$ , sea  $B \in \mathcal{A}$  y  $0 < P[B] < 1$ , entonces  $\forall A \in \mathcal{A}$

$$P[A] = P[A/B] P[B] + P[A/B^c] P[B^c]$$

### Nota.

El Teorema de la Probabilidad Total se cumple también si  $n = \infty$ .

El teorema de la probabilidad total y su corolario son de particular utilidad para aquellos experimentos que tienen fases o etapas, es decir, el experimento consiste primero en ejecutar una cosa (primera etapa) y después otra cosa (segunda etapa). El ejemplo 28 es de este tipo de experimentos; en la etapa 1 se selecciona a una urna y en la etapa 2 se extrae una bola de la urna seleccionada. En tales experimentos, si  $B_j$  es un evento definido exclusivamente en términos de la primera etapa y  $A$  es otro evento definido en términos de la segunda etapa, entonces será sencillo encontrar  $P[B_j]$  y  $P[A/B_j]$  para  $j = 1, 2, \dots, n$ . Recuerde que en los experimentos de etapas, es más cómodo trabajar con eventos condicionados por resultados de la primera etapa.

### Teorema 31, Fórmula de Bayes.

Para un espacio probabilístico dado  $(\Omega, \mathcal{A}, P[\cdot])$ , si  $B_1, B_2, \dots, B_n$  es una partición del espacio muestral, entonces se tiene que para cualquier  $A \in \mathcal{A}$

$$P[B_k/A] = \frac{P[A/B_k] \cdot P[B_k]}{\sum_{j=1}^n P[A/B_j] \cdot P[B_j]}$$

### Demostración

Por la definición de probabilidad condicional y por el teorema de la probabilidad total se tiene que

$$P[B_k/A] = \frac{P[B_k \cap A]}{P[A]} = \frac{P[A/B_k]P[B_k]}{\sum_{j=1}^n P[A/B_j]P[B_j]}$$

### Corolario.

Sea  $(\Omega, \mathcal{A}, P[\cdot])$  un espacio probabilístico dado y sean  $A, B \in \mathcal{A}$   $P[A] > 0$  y  $0 < P[B] < 1$ , entonces

$$P[B/A] = \frac{P[A/B] \cdot P[B]}{P[A/B] \cdot P[B] + P[A/B^c] \cdot P[B^c]}$$

### Nota.

La Fórmula de Bayes y su corolario son aún válidos para si  $n = \infty$ .

Al igual que el caso de la Probabilidad Total, la Fórmula de Bayes es particularmente útil para aquellos experimentos que consisten de fases o etapas. Si  $B_j$  con  $j = 1, \dots, n$ , es un evento definido exclusivamente en términos de la primera etapa y  $A$  es otro evento esta vez definido en todo el experimento incluyendo la segunda etapa, utilizando la regla de Bayes la pregunta  $P[B_k/A]$  es expresada en sentido inverso, es decir  $P[A/B_k]$ . Esto es útil en casos en que por las condiciones normales del experimento, pueda resolverse o sea más sencillo resolver  $P[A/B_k]$  en vez de  $P[B_k/A]$ .

### Teorema 32, Regla del Producto.

Sean  $(\Omega, \mathcal{A}, P[\cdot])$  un espacio probabilístico dado y  $A_1, A_2, \dots, A_n \in \mathcal{A}$  tales que  $P[A_1 \cap A_2 \cap \dots \cap A_n] > 0$ , entonces

$$P[A_1 \cap A_2 \cap \dots \cap A_n] = P[A_1]P[A_2/A_1]P[A_3/A_1 \cap A_2] \dots P[A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1}].$$

### Demostración

La demostración puede ser obtenida empleando inducción matemática y se deja a modo de ejercicio.

Al igual que los dos anteriores teoremas, la regla del producto es de utilidad para aquellos experimentos definidos en varias etapas. Suponga que el experimento consta de  $n$  etapas y  $A_j$  es un evento denotado en términos de la  $j$ -ésima etapa del experimento. Entonces  $P[A_j/A_1 \cap A_2 \cap \dots \cap A_{j-1}]$  es la probabilidad condicional de un evento descrito en términos de la  $j$ -ésima etapa condicionado por lo que ocurra en las etapas  $1, 2, \dots, j-1$ . La regla del producto expresa a

$P[A_1 \cap A_2 \cap \dots \cap A_n]$  en términos de las probabilidades condicionales más resolubles  $P[A_j/A_1 \cap A_2 \cap \dots \cap A_{j-1}]$  para  $j = 2, \dots, n$ .

### Ejemplo 28.

Se tienen 5 urnas numeradas de la 1 a la 5. Cada  $i$ -ésima urna contiene 10 bolas, de las cuales exactamente  $i$  son defectuosas y  $10 - i$  no son defectuosas. Por ejemplo, la urna 3 tiene 3 bolas defectuosas y 7 bolas no defectuosas. El experimento aleatorio consiste en 2 pasos: primero seleccionar una urna al azar y después extraer una bola de la urna seleccionada (el experimentador no conoce cual urna fue seleccionada). Contestemos las siguientes preguntas: (i) ¿cuál es la probabilidad de que salga una bola defectuosa? y (ii) si se ha extraído una bola defectuosa, ¿cual es la probabilidad de que esta provenga de la urna 5?

### Solución.

Sea el evento  $A = \{\text{se extraiga de la urna una bola defectuosa}\}$  y  $B_i = \{\text{la urna } i\text{-ésima es seleccionada con } i=1, \dots, 5\}$ . Tenemos entonces que  $P[B_i] = 1/5$  con  $i = 1, \dots, 5$  y  $P[A/B_i] = i/10$  con  $i = 1, \dots, 5$ . Entonces la pregunta (i) es precisamente  $P[A]$ . Usando el teorema de la probabilidad total tenemos que

$$P[A] = \sum_{i=1}^5 P[A/B_i]P[B_i] = \sum_{i=1}^5 \frac{i}{10} \cdot \frac{1}{5} = \frac{1}{50} \sum_{i=1}^5 i = \frac{1}{50} \cdot \frac{5 \cdot 6}{2} = \frac{3}{10}.$$

Observe que en total hay 50 bolas de las cuales 15 son defectuosas y sin embargo la probabilidad de que una bola salga defectuosa no es  $15/50$ .

Por otro lado, la pregunta (ii) esta dada por  $P[B_5/A]$ . Dado que la urna 5 tiene más bolas defectuosas que cualquiera otro, esperamos que  $P[B_5/A] > P[B_i/A]$  para  $i = 1, \dots, 4$ . De hecho esperamos que  $P[B_5/A] > P[B_4/A] > \dots > P[B_1/A]$ . Usando la fórmula de Bayes tenemos que

$$P[B_5/A] = \frac{P[A/B_5]P[B_5]}{\sum_{i=1}^5 P[A/B_i]P[B_i]} = \frac{\frac{1}{2} \cdot \frac{1}{5}}{\frac{3}{10}} = \frac{1}{3}.$$

De manera similar

$$P[B_k/A] = \frac{\frac{k}{10} \cdot \frac{1}{5}}{\frac{3}{10}} = \frac{k}{15}, \text{ con } k = 1, \dots, 5 \text{ lo cual comprueba nuestras}$$

sospechas. Note usted que los eventos  $B_1, \dots, B_5$  son equiprobables, sin embargo las probabilidades condicionales  $P[B_k/A]$  no lo son. Observe también que

$$\sum_{k=1}^5 P[B_k / A] = \sum_{k=1}^5 \frac{k}{15} = \frac{1}{15} \sum_{k=1}^5 k = \frac{1}{15} \cdot \frac{5 \cdot 6}{2} = 1.$$

### Ejemplo 29.

Un estudiante esta contestando un examen de opción múltiple. En una pregunta dada, el estudiante no conoce la respuesta, lo que puede resultar en que conteste correctamente o que conteste incorrectamente al intentar adivinar la respuesta adecuada. El examen consta de 5 opciones por pregunta como es costumbre. El encargado de calificar el examen se confronta entonces al siguiente dilema, habiendo contestado correctamente una pregunta, quisiera saber cual es la probabilidad de que el estudiante sabia la respuesta? Sea  $p$  la probabilidad de que el estudiante conozca la respuesta y  $1-p$  la probabilidad de que el estudiante adivine la respuesta. Asumamos que la probabilidad de que el alumno escoja la respuesta correcta dado que esta adivinando es igual a  $1/5$ . (Esta suposición pudiera no ser verdadera, ya que el alumno aun que no conozca la respuesta correcta, si pudiera saber que algunas respuestas son incorrectas con lo que debiera ser mayor a  $1/5$ ). Sea el evento  $A = \{\text{el alumno contesto adecuadamente}\}$  y el evento  $B = \{\text{el alumno conocía la respuesta correcta}\}$ , ¿cuánto vale  $P(B/A)$ ? Usando el corolario de la fórmula de Bayes, tenemos que

$$P[B/A] = \frac{P[A/B]P[B]}{P[A/B]P[B] + P[A/B^c]P[B^c]} = \frac{1 \cdot p}{1 \cdot p + \frac{1}{5}(1-p)}.$$

$$\text{Note que } \frac{p}{p + \frac{1}{5}(1-p)} \geq p.$$

### Ejemplo 30.

Una urna contiene 10 bolas de las cuales 3 son negras y 7 son blancas. El siguiente juego es realizado: en cada ensayo una bola es seleccionada al azar, se registra su color, y es reemplazada en la urna junto con dos bolas más del mismo color. ¿Cuál es la probabilidad de que una bola negra sea seleccionada en los primeros tres ensayos? Sea el evento  $B_j = \{\text{bola negra es seleccionada en la } j\text{-ésima extracción}\}$ , estamos buscando  $P[B_1 \cap B_2 \cap B_3]$ . Por la regla del producto

$$P[B_1 \cap B_2 \cap B_3] = P[B_1]P[B_2/B_1]P[B_3/B_1 \cap B_2] = \frac{3}{10} \cdot \frac{5}{12} \cdot \frac{7}{14} = \frac{1}{16}.$$

### Ejemplo 31.

Suponga que una urna contiene  $M$  bolas de las cuales  $K$  son negras y  $M - K$  son blancas. Se extrae una muestra de tamaño  $n$ . Encuentre la probabilidad de

que la  $j$ -ésima bola seleccionada sea negra dado que la muestra contiene exactamente  $r$  bolas negras? Intuitivamente esperamos que la respuesta sea  $r/n$ . Encuentre la respuesta considerando (i) muestreo con reemplazo y (ii) muestreo sin reemplazo

### Solución.

Sea el evento  $A_r = \{\text{la muestra contiene exactamente } r \text{ bolas negras}\}$  y  $B_j = \{\text{la } j\text{-ésima bola seleccionada es negra}\}$ . Entonces buscamos  $P[B_j/A_r]$ . Consideremos primero (i):

$$P[A_r] = \binom{n}{r} \frac{K^r (M-K)^{n-r}}{M^n} \text{ y } P[A_r/B_j] = \binom{n-1}{r-1} \frac{K^{r-1} (M-K)^{n-r}}{M^{n-1}} \text{ por (AN4).}$$
 Ya que las bolas son reemplazadas,  $P[B_j] = K/M$  para toda  $1 \leq j \leq n$ , entonces

$$P[B_j/A_r] = \frac{P[A_r/B_j]P[B_j]}{P[A_r]} = \frac{\binom{n-1}{r-1} \frac{K^{r-1} (M-K)^{n-r}}{M^{n-1}} \cdot \frac{K}{M}}{\binom{n}{r} \frac{K^r (M-K)^{n-r}}{M^n}} = \frac{r}{n}.$$

Ahora consideremos (ii):

$$P[A_r] = \frac{\binom{K}{r} \binom{M-K}{n-r}}{\binom{M}{n}} \text{ y } P[A_r/B_j] = \frac{\binom{K-1}{r-1} \binom{M-K}{n-r}}{\binom{M-1}{n-1}} \text{ por (AN6).}$$

Si  $P[B_j] = \sum_{i=0}^{j-1} P[B_j/C_i]P[C_i]$  donde el evento  $C_i = \{\text{extraer exactamente } i \text{ bolas negras en las primeras } j-1 \text{ extracciones}\}$ . Note que

$$P[C_i] = \frac{\binom{K}{i} \binom{M-K}{j-1-i}}{\binom{M}{j-1}} \text{ y } P[B_j/C_i] = \frac{K-i}{M-j+1}, \text{ entonces}$$

$$P[B_j] = \sum_{i=1}^{j-1} \frac{K-i}{M-j+1} \cdot \frac{\binom{K}{i} \binom{M-K}{j-i-1}}{\binom{M}{j-1}} = \frac{K}{M}. \text{ Finalmente,}$$

$$P[B_i/A_r] = \frac{P[A_r/B_i]P[B_i]}{P[A_r]} = \frac{\left[ \frac{\binom{K-1}{r-1} \binom{M-K}{n-r}}{\binom{M-1}{n-1}} \right] \frac{K}{M}}{\frac{\binom{K}{r} \binom{M-K}{n-r}}{\binom{M}{n}}} = \frac{r}{n}.$$

Se ha obtenido el resultado esperado con reemplazo o sin reemplazo.

### Independencia de Eventos.

Si  $P[A/B]$  no depende del evento  $B$ , esto es, si  $P[A/B] = P[A]$ , parece natural decir que el evento  $A$  es independiente del evento  $B$ . En otras palabras, dos eventos son independientes, si la ocurrencia de uno no implica la ocurrencia o no ocurrencia del otro evento. Esta idea da lugar a la siguiente definición.

#### Definición 20, Eventos Independientes.

Para un espacio probabilístico dado  $(\Omega, \mathcal{A}, P[\cdot])$  y  $A, B \in \mathcal{A}$ . Se dice que los eventos  $A$  y  $B$  son *independientes* si y solo si alguna de las siguientes condiciones se satisface:

- (i)  $P[A \cap B] = P[A]P[B]$ .
- (ii)  $P[A/B] = P[A]$  si  $P[B] > 0$ .
- (iii)  $P[B/A] = P[B]$  si  $P[A] > 0$

#### Nota.

Algunos autores usan el término "Estocasticamente Independientes" en vez de "Independientes".

#### Ejemplo 32.

Considere el experimento de lanzar dos dados. Sea los eventos  $A = \{\text{los dados suman número par}\}$ ,  $B = \{\text{sale comodín en el primer dado}\}$  y  $C = \{\text{los dados suman 7}\}$ . Contestar las siguientes preguntas,

- (i) ¿ $A$  y  $B$  son Independientes?
- (ii) ¿ $A$  y  $C$  son Independientes?
- (iii) ¿ $B$  y  $C$  son Independientes?

Tenemos que  $P[A/B] = 1/2 = P[A]$ , por lo que A y B sí son independientes. Por otro lado,  $P[A/C] = 1 \neq 1/2 = P[A]$ , por lo que A y C no son independientes. Finalmente  $P[C/B] = 1/6 = P[C]$ , entonces B y C son también independientes.

La Independencia entre eventos y eventos Mutuamente Excluyentes son dos propiedades muy diferentes, aunque estas sí existe relación entre ambas. Por ejemplo, dos eventos mutuamente excluyentes A y B son además independientes si y solo si  $P[A]P[B] = 0$ , lo cual es verdadero si alguno de los dos, A o B, tiene probabilidad igual a cero. Si  $P[A] \neq 0$  y  $P[B] \neq 0$ , entonces la independencia de A y B implica que no son mutuamente excluyentes, y A y B mutuamente excluyentes implica que no son independientes. Sin embargo, la independencia de A y B implica a veces la Independencia de otros eventos:

### Teorema 33.

Sean A y B dos eventos independientes definidos en un espacio probabilístico dado  $(\Omega, \mathcal{A}, P[\cdot])$ , entonces A y  $B^c$  son independientes,  $A^c$  y B son independientes y  $A^c$  y  $B^c$  son independientes.

### Demostración.

$$P[A \cap B^c] = P[A] - P[A \cap B] = P[A] - P[A]P[B] = P[A](1 - P[B]) = P[A]P[B^c]$$

De manera similar se demuestran los otros dos casos.

La noción de independencia se puede extender a más de dos eventos:

### Definición 21, Eventos Independientes.

Para un espacio probabilístico dado  $(\Omega, \mathcal{A}, P[\cdot])$ , sean  $A_1, A_2, \dots, A_n \in \mathcal{A}$ . Se dice que los eventos  $A_1, A_2, \dots, A_n$  son independientes si y solo si

$$P[A_i \cap A_j] = P[A_i] P[A_j], \text{ para } i \neq j \text{ y}$$

$$P[A_i \cap A_j \cap A_k] = P[A_i] P[A_j] P[A_k], \text{ para } i \neq j, j \neq k, i \neq k \text{ y}$$

$$\vdots$$

$$P\left[\bigcap_{i=1}^n A_i\right] = \prod_{i=1}^n P[A_i].$$

Lo primero que uno se pregunta, es si todas las condiciones de la definición son requeridas. Por ejemplo,  $P[A_1 \cap A_2 \cap A_3] = P[A_1] P[A_2] P[A_3]$  implica que  $P[A_1 \cap A_2] = P[A_1] P[A_2]$ ? Obviamente no por que si  $P[A_3] = 0$ , se cumple que  $P[A_1 \cap A_2 \cap A_3] = P[A_1] P[A_2] P[A_3] = 0$  pero  $P[A_1 \cap A_2] \neq P[A_1] P[A_2]$  si  $A_1$  y  $A_2$  no

son independientes. Pudiera uno también preguntarse si la independencia por parejas implica independencia? De nuevo la respuesta es negativa.

## Bibliografía.

- Alexander M. Mood, Franklin A. Graybill, Duane C. Boes  
Introduction to the Theory of Statistics, Third Edition  
USA, McGraw-Hill, 1974
- Chester Wallance Jordan Jr.  
Life Contingences, Second Edition  
Chicago USA, The Society of Actuaries, 1975
- Newton L. Bowers, Jr., Hans U. Gerber, James C. Hickman, Donald A. Jones,  
Cecil J. Nesbitt  
Actuarial Mathematics  
Chicago USA, The Society of Actuaries, 1986
- George C. Canavos  
Probabilidad y Estadística. Aplicaciones y Métodos.  
España, McGraw-Hill, 1998
- Shoichiro Nakamura  
Análisis Numérico y Visualización Gráfica con Matlab®  
México, Prentice-Hall, 1997
- William Wooton, Edwin F. Beckenbach, Frank F. Fleming  
Geometría Analítica Moderna  
México D.F., Publicaciones Cultural S.A., 1977
- Martin Braun  
Ecuaciones Diferenciales y sus Aplicaciones.  
México, Iberoamericana, 1990
- AMA, AMIS  
Tabla de Mortalidad México 2000.  
México D.F., Comité de Mortalidad AMA-AMIS, Junio de 2000.
- Mario Gabriel Taba Herlindo  
Estudio de la Aplicación de una Nueva Tabla de Mortalidad en el  
Mercado Mexicano de Seguros  
México D.F., Tesis Facultad de Ciencias UNAM, 1973
-

**María Desideree Gorostieta García**

Tabla de Mortalidad Experiencia Mexicana (62-67). Análisis y Ajuste por medio de las Funciones de Levi.  
México D.F., Tesis Facultad de Ciencias UNAM, 1990

**Cecilia Meneses Molina**

Proyecto para la Elaboración de una Tabla de Mortalidad de los Asegurados en Instituciones Mexicanas de Seguros  
México D.F., Tesis Facultad de Ciencias UNAM, 1964

**Diario Oficial de la Federación**

Primera Sección del Diario Oficial del Viernes 31 de Diciembre de 1999  
México D.F.

**Manuel Mendoza, Ana María Madrigal, Evangelina Martínez**

Análisis Exploratorio de la Información del Seguro de Vida en México  
1991-1998  
México D.F., Serie de Trabajo No. 76 de la CNSF, Octubre 1999

**Manuel Mendoza, Ana María Madrigal, Evangelina Martínez**

Modelos Estadísticos de Mortalidad. Análisis de Datos 1991-1998  
México D.F., Serie de Trabajo No. 77 de la CNSF, Diciembre 1999

**Miguel Ángel García**

Notas en Clase del Curso Probabilidad I  
México D.F., Facultad de Ciencias, 1996