PROBABILIDAD LÓGICA

Alberto A. Alonso

Ediciones anticipar

PROBABILIDAD LÓGICA

“Es remarcable que una ciencia la cual comenzó con el estudio sobre las chances en juegos de azar, se haya convertido en el objeto más importante del conocimiento humano… las preguntas más importantes sobre la vida son, en su mayor parte, en realidad sólo problemas de probabilidad”

Pierre Simon, Marqués de Laplace (1749-1827)

Toda obra grande, en arte como en ciencia,
es una gran pasión al servicio de una gran idea.
Santiago Ramón y Cajal

PROBABILIDAD LÓGICA

Alberto A. Alonso

Ediciones anticipar

Buenos Aires - Argentina

Alberto A. Alonso es Ingeniero Químico por la Universidad Nacional de La Plata, posee una certificación internacional en Administración de Riesgos por ALARYS y posee, también, un curso de especialización en Estadística Descriptiva por el CONICET. En su vida profesional, ha sido declarado "Experto en temas de Ingeniería" por el Ministerio de Educación y Justicia de la Nación -Resolución D.N.A.U. Nº 86 del año 1987. Actualmente es Profesor titular de Estadística Aplicada en el IUPFA, para las Licenciaturas en Seguridad, Accidentología y Prevención Vial, Trabajo Social y para la carrera de Ingeniería en Siniestros y Seguridad Ambiental.

Ver CV completo en: http:// www.anticiparconsultoria.com

Alonso, Alberto A.
Probabilidad lógica. - 1a ed. - Ciudad Autónoma de Buenos Aires: Anticipar, 2014.
E-Book.

ISBN 978-987-45197-8-8

1. Probabilidades. 2. Lógica. I. Título
CDD 519.2

Fecha de catalogación: 30/05/2014

Ediciones anticipar: http://anticiparconsultoria.com

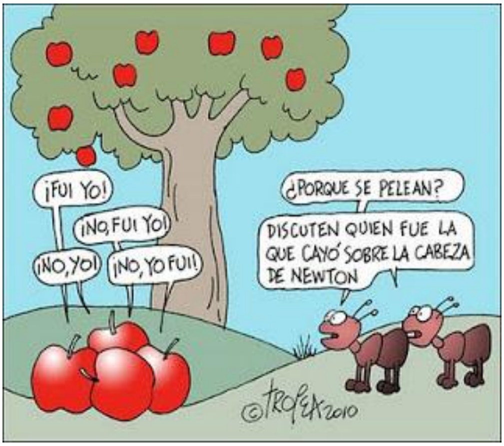
Esmeralda 582 – Piso 8º Of. 30 – (C1007ABD) – Ciudad Autónoma de Buenos Aires
Primera edición. Marzo 2014.

Si te atreves a enseñar,
no dejes de aprender.
John Cotton Dana

PREFACIO

Cuando decidí que el título de esta obra fuese Probabilidad Lógica, mi intensión fue, quizás, siguiendo los ideales de Schlick, Carnap, Neurath y tantos otros filósofos y científicos europeos, principalmente alemanes, que a principios del Siglo XX fundaron el *Círculo de Viena*, que los estudiantes al aprender probabilidades en vez de pensar en las matemáticas puras pensasen en la lógica de las ciencias. Por eso, el libro fue planteado utilizando un lenguaje coloquial antes que esos lenguajes algo abstractos utilizados en el ámbito cerrado de algunos ambientes científicos.

Y esto creo que es así, porque a los grandes descubrimientos se llegó por la observación y la deducción y luego se les dio a esos hallazgos un formato matemático, conocido como ley, que permitía su cálculo y reproducción. Seguramente, Isaac Newton no escribió: $F = g.\frac{m_1 \times m_2}{r^2}$ investigando la física y la matemática y, a partir de ello, después vio caer la manzana[1]. El proceso fue justamente al revés. No en vano la observación es uno de los cuatro métodos que conforman la recogida de información.



La *probabilidad a priori* nace antes que la *probabilidad a posteriori*, y cuando el gran médico y matemático italiano Gerolamo Cardano[2], un gran transgresor en su época, allá por el año 1560 escribió su *liber de ludo aleae (Libro de los juegos de azar)*, que fue la primera obra importante relacionada con el cálculo de probabilidades en los juegos de azar, seguramente habrá efectuado miles y miles

---

[1] Imagen del chiste: http://blogfisicayquimicaelena.blogspot.com.ar/2011/05/existio-la-manzana-de-newton.html

[2] Fuente de la imagen: http://smk1-girolamo.blogspot.com.ar/2012/03/frases-celebres-de-cardamo.html

de observaciones con dados y monedas antes de analizar matemáticamente la probabilidad buscando que, de alguna manera, la fórmula matemática reprodujese la realidad tocástica.

En la probabilidad lógica o la probabilidad vista desde la lógica, existen ciertos enunciados que describen resultados observacionales que se traducen en una hipótesis a ser confirmada o no. Esta relación tiene la forma de un razonamiento inductivo con premisas singulares y una conclusión general.

Siempre les digo a mis alumnos que la universidad no me ha contratado para dar clases. Lo ha hecho para que enseñe y la enseñanza no debe ser autoritaria. El docente no está para inculcar su verdad. La función del profesor es la de incitar a los alumnos a ver y percibir lo que se le enseña. La enseñanza debe abrir las puertas al debate y a la imaginación para que finalmente triunfe la deducción y la predicción.

La probabilidad, tan ligada a la estadística y a la ciencia actuarial, sin duda usa, y mucho, a las matemáticas puras y sus desarrollos numéricos no pueden hacer otra cosa que infundir cierto miedo y respeto a los legos. Sin embargo, la probabilidad como herramienta básica de todas las profesiones, puede ser entendida y aplicada utilizando cálculos muy básicos que no exceden las operaciones matemáticas aprendidas en el colegio primario y secundario.

Muchas actividades se desarrollaron a partir del impulso exhaustivo de la estadística y la probabilidad, como lo fue la técnica del seguro. Pero, la probabilidad básica, forma parte de las cosas de todos los días.

De allí que la estadística y la probabilidad hoy forman parte del currículo de cualquier carrera, pertenezca ésta a las ciencias duras o a las del comportamiento.

Desde la didáctica, creemos que hay cuatro características que conviven en este libro.

Primera: La naturalidad o llaneza. La redacción tiende a ser algo irreverente con respecto al cientificismo puro. Quiere ser real, directa y carente de solemnidad.

Segunda: El formato. El texto está bastante complementado con gráficos e ilustraciones que ayuda a comprender y asimilar la letra escrita.

Tercera: La ilustración. Cada científico que mencionamos en la obra ha sido acompañado de una imagen. Creemos que resulta muy afectivo y efectivo que el lector “vea” como era la persona que tanto ayudó y se esforzó por desarrollar la ciencia de la probabilidad.

Cuarta: La ejercitación. Cada tema teórico es acompañado de ejemplos y al final de la unidad por ejercicios resueltos, para que el lector comprenda para qué sirve

y en que se utiliza lo que está leyendo.
Para finalizar, solo deseamos decirle al lector o a la lectora una sola cosa: Que lo disfrute.
Ojalá lo logremos...

Alberto Adriano Alonso
La Plata, invierno de 2014

http://www.anticiparconsultoria.com

alonso@anticiparconsultoria.com

Agradecimiento

Les dedico esta obra a todos los que me quieren y confían en mí.

Reconocimiento

Aunque no lo conozca personalmente, deseo reconocer muy especialmente a Adrian Paenza[3], por coincidir totalmente con él, en el sentido de que las ciencias deben enseñarse de manera sencilla con el fin que el estudiante las entienda y también de una manera amena para que, además, comprenda su utilidad y aplicación.

---

[3] **Adrián Arnoldo Paenza** (n. Buenos Aires, 9 de mayo de 1949) es un periodista y doctor en ciencias matemáticas por la Facultad de Ciencias Exactas y Naturales (UBA). Así describe su carrera uno de sus últimos libros (2010. *Matemática... ¿Estás ahí? La vuelta al mundo en 34 problemas y 8 historias*): "Nació en Buenos Aires en 1949. Es doctor en Matemáticas por la Universidad de Buenos Aires, donde se desempeña actualmente como profesor asociado del Departamento de Matemática de la Facultad de Ciencias Exactas y Naturales. Es, además, periodista. Actualmente es columnista especial de Página/12. Publicó en esta misma colección [*Ciencia que ladra*] los cuatro tomos iniciales de la serie *Matemática… ¿estás ahí?*, que han sido un éxito de ventas en la Argentina, en otros países de Latinoamérica y también en Alemania y España, donde se han editado los dos primeros episodios. Asimismo, sus libros han sido publicados (o lo serán próximamente) en Rusia, Italia, República Checa, Brasil y Portugal. En 2007 recibió el premio Konex de platino en el rubro “Divulgación científica”." En 2014 recibió el Premio Lilavati del ICM por su labor en la divulgación de las matemáticas. Fuente: extractado de: http://es.wikipedia.org/wiki/Adri%C3%A1n_Paenza

ÍNDICE

¿Cómo osamos hablar de leyes del azar?
¿No es, acaso, el azar la antítesis de cualquier ley?
Bertrand Russell[4]

UNIDAD 1

MUESTREO[5]

Antes de comenzar a ver el tema de probabilidades, tenemos que analizar algunos conceptos de muestreo que van muy de la mano con el tratamiento de la probabilidad.

[4] Durante los 98 años de vida de Bertrand Russell (1872-1970), este fue testigo de cambios en la sociedad y en la política internacional que nadie hubiese podido imaginar en la Inglaterra victoriana en la que creció y se educó: Dos guerras mundiales, la lucha por el sufragio femenino, la crisis económica, el desarrollo del capitalismo, el comunismo y el fascismo, la lucha por los derechos civiles en EE.UU., la guerra de Vietnam... Repasar la vida de Bertrand Russell supone hacer un recorrido histórico no solo por las personalidades del mundo de la Filosofía con quienes se relacionó (Mc. Taggart, G. E. Moore, Whitehead, Ludwig Wittgenstein...), sino también por los acontecimientos mencionados y por muchas otras figuras relevantes con las que mantuvo algún tipo de contacto (Keynes, Lenin o Joseph Conrad, entre otros). Fuente: http://www.filosofos.net/russell/russell_bio.htm

[5] Fuente de la imagen: http://dospensamiento.blogspot.com.ar/

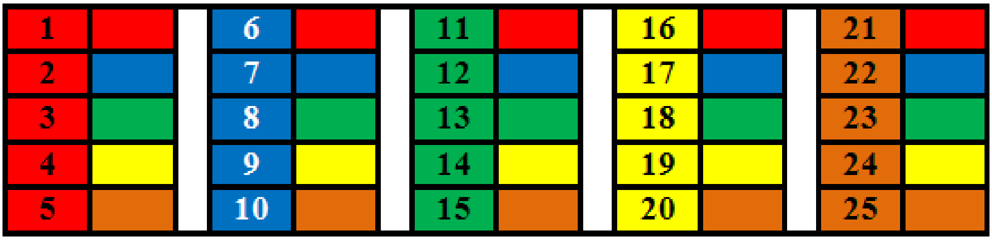
Se entiende por muestreo al procedimiento utilizado para obtener una muestra, por lo que la muestra no es otra cosa que el resultado de un muestreo.
Ya habíamos visto, en repetidas ocasiones, algo que es como el *abc* de la Estadística inferencial, y está referido a que, para poder generalizar válidamente las características de una muestra a la población de la cual proviene, tanto en los experimentos de prueba de hipótesis como en los de estimación de parámetros, la muestra debe ser aleatoria.
En estadística, una hipótesis es un supuesto que no ha sido verificado pero que podría ser el resultado o posible solución de una hipótesis[6] que el experimentador desea comprobar.
Por su parte, la estimación de parámetros es el conjunto de técnicas que permiten asignar un valor al parámetro de una población a partir del valor del estadístico de una muestra.
Para que el muestreo sea probabilístico o aleatorio, es requisito que:

- Todos y cada uno de los elementos de la población tengan la misma probabilidad de ser seleccionados.
- Que todas y cada una de las muestras tengan la misma probabilidad de ser elegidas.

Procediendo de esta manera, la muestra que sea elegida deberá ser considerada una muestra aleatoria.
Supongamos que tenemos cinco tejidos de colores distintos: rojo, azul, verde, amarillo y terracota. Quiere decir que, en este caso, la población consta de tan solo cinco [7]individuos o elementos. Un modisto tiene que elegir dos de esos colores para hacer un traje *sport* de dama, con la condición de que el primer color del par elegido corresponderá a la chaqueta y el segundo, a la falda.
Por lo tanto, tenemos que elegir, *aleatoriamente*, una muestra de tamaño n=2. La cantidad total de muestras de tamaño n= 2 (o que contienen dos individuos) que podemos lograr a partir de una población de tan solo 5 elementos, son 25 y son las que se pueden observar en las cinco columnas de la Figura 1.1.

[6] Hipótesis: Científicamente, enunciados teóricos supuestos, no verificados pero probables referentes a variables o relaciones entre variables. En relación al problema de investigación, soluciones probables, previamente seleccionadas, al problema planteado que el científico propone para comprobar, a través de todo el proceso de investigación, si son confirmadas por los hechos. SIERRA BRAVO, Restituto. (1991). *Diccionario Práctico de Estadística*. Madrid, Ed. Paraninfo.
[7] Cada uno de los elementos de la población. Fuente: http://recursostic.educacion.es/secundaria/edad/2esomatematicas/2quincena12/2quincena12_contenidos_1a.htm

Figura 1.1. Vista de las 25 muestras de dos colores cada una



Hemos formado 25 combinaciones o muestras entre los cinco colores de la población que incluyen la totalidad de muestras posibles. Aquí se cumple el requisito de que todos y cada uno de los elementos de la población tuvieron la misma probabilidad de ser seleccionados.
A continuación, tenemos que elegir una única muestra de entre las 25 muestras disponibles. Una posibilidad es utilizar un bolillero con 25 bolillas y sacar una de ella. Supongamos que ese número fue el 15, por lo tanto, la muestra *verde-terracota* habrá sido la seleccionada. En estas condiciones, el modisto realizará el traje de dama con una chaqueta color verde y una falda color terracota.
Aquí se ha cumplido el segundo requisito, referido a que todas las muestras han tenido la misma probabilidad de ser elegidas.

MUESTREO ALEATORIO[8]

Como habíamos visto recientemente, para poder comprobar una hipótesis o estimar un parámetro, la muestra con la cual trabajaremos no puede ser cualquier subconjunto de la población. Es imprescindible que la misma sea aleatoria.

Vemos, entonces, que si el proceso garantiza que cada muestra posible de un tamaño dado tenga la misma probabilidad de ser elegida, entonces también garantiza que todos los miembros de la población tengan la misma posibilidad de ser elegidos en la muestra.

Para ilustrar la definición, consideremos la situación en la que tenemos una población de seis individuos (N= 6)[9] que numeramos del uno al seis, ellos serian:

---

8 Fuente de la imagen: http://javierchavez07.blogspot.com.ar/2013/07/ejercicio-de-muestreo-aleatorio.html

9 N. del A. La cantidad de individuos de una población se designa con la letra **N** (mayúscula), mientras que la cantidad de individuos de una muestra se designa con la letra **n** (minúscula).

1	2	3	4	5	6

Ahora queremos extraer, de manera aleatoria, una muestra de tamaño n=2, es decir una muestra de dos individuos que surja de la población que consta de seis individuos. Tengamos en cuenta que, por lo general, una población tendrá más individuos.
La hemos restringido a tan solo seis individuos para facilitar la comprensión de lo que queremos explicar. Supongamos que estamos realizando un muestreo de esta población y que tomamos un individuo a la vez, para luego regresarlo con el resto de los individuos antes de extraer otro. Esto se llama muestreo con reemplazo y se analiza posteriormente en esta misma unidad.
En la tabla 1.1. mostramos todas las muestras posibles, de tamaño n = 2, que podemos obtener de la población, mediante este método de muestreo:

Tabla 1.1.					
1-1	2-1	3-1	4-1	5-1	6-1
1-2	2-2	3-2	4-2	5-2	6-2
1-3	2-3	3-3	4-3	5-3	6-3
1-4	2-4	3-4	4-4	5-4	6-4
1-5	2-5	3-5	4-5	5-5	6-5
1-6	2-6	3-6	4-6	5-6	6-6

Vemos que son 36 muestras de tamaño n=2, las que obtenemos al realizar el muestreo con reemplazo. Para producir un muestreo aleatorio, el proceso debe ser tal que:

- tanto las 36 muestras,
- como todos los individuos de la población (1,2, 3, 4, 5 y 6) tengan la misma probabilidad de ser seleccionados en la muestra.

La muestra debe ser aleatoria por dos razones. La primera, es que para generalizar una característica que presenta una muestra a toda la población, es necesario aplicarle las leyes de la probabilidad. Si la muestra no se ha generado mediante un proceso que garantice que cada muestra posible del tamaño dado tenga la misma posibilidad de ser escogida, no podremos aplicar las leyes de la probabilidad.

La segunda razón para la utilización del muestreo aleatorio, es qué para asegurar la generalización de una característica de la muestra, a toda la población, es necesario que esta muestra sea representativa de la población en cuestión.
Una forma de lograr tal representatividad, es elegir la muestra mediante un proceso que garantice que todos los miembros de la población tengan la misma probabilidad de ser escogidos. Así, requerir que una muestra sea aleatoria permite utilizar las leyes de la probabilidad sobre la muestra misma y, con ella, que sea representativa de la población.
Es tentador pensar que podríamos lograr la representatividad de una muestra mediante métodos distintos al del muestreo aleatorio. Sin embargo, con mucha frecuencia el procedimiento utilizado produce una muestra sesgada (no representativa). Un ejemplo de esto es la famosa encuesta *Literary Digest* [10], la cual predijo una victoria aplastante de Landon (57 a 43%) sobre Roosvelt, en las elecciones presidenciales de los EE.UU del año 1936. En la realidad, Roosevelt ganó la elección con el 62% de los votos. ¿Fue una falla de la estadística? No. La predicción del *Literary Digest* fue un error garrafal de muestreo. ¿Por qué? Un análisis posterior reveló que el error se debió a que la muestra no era representativa de la población votante, sino que era muy sesgada. Los individuos fueron elegidos de fuentes como el directorio telefónico, registros de clubes y de propietarios de automóviles. Estos registros excluyeron de manera sistemática a los pobres, quienes tal vez no tenían teléfonos o automóviles. Lo que ocurrió, fue que los pobres votaron en forma aplastante por Roosevelt y la predicción no los había tenido en cuenta como votantes. Ver en el

TÉCNICAS PARA EL MUESTREO ALEATORIO

Un estudio profundo de las formas mediante las cuales generaríamos muestras aleatorias, está más allá del objetivo de esta unidad. Este tema puede ser complejo, en particular al trabajar con encuestas. Sin embargo, presentaremos algunas técnicas de uso común, junto con algunas situaciones sencillas, de modo que podamos tener una idea de lo que esto implica.

Técnicas populares
Continúa ...

[10] Una buena ilustración de la encuesta, puede leerse en el artículo de Adrián Paenza: http://www.pagina12.com.ar/diario/contratapa/13-211281-2013-01-06.html

La probabilidad de que los semáforos nos den luz roja, es directamente proporcional al apuro que llevamos

Anónimo

UNIDAD 2

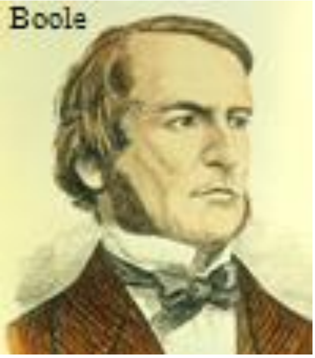
PROBABILIDAD[11]

Una parte de la Estadística Inferencial está ligada al cálculo de las probabilidades, un área asombrosa y sorprendente que nos enseña que el azar es todo.
Los temas de muestreo aleatorio y de probabilidad son fundamentales para la metodología de la Estadística Inferencial.

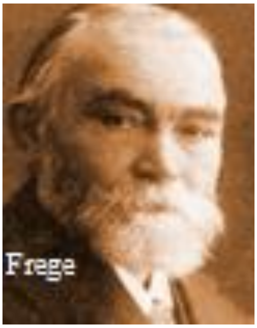
PROBABILIDAD, MATEMÁTICAS Y LÓGICA

En el último cuarto del siglo XIX se vivió un episodio apasionante en la historia de las matemáticas, que la ligaría desde entonces a la historia de la lógica.
Primero, George Boole[12], un matemático británico que vivió entre 1815 y 1864

[11] Fuente de la imagen: http://nuneznjaimer.mex.tl/frameset.php?url=/
[12] Fuente de la imagen: http://wintablet.info/2012/06/gracias-george/



en su libro *Mathematical Analysis of Logic* (Análisis Matemático de la Lógica) trató de presentar la lógica como parte de las matemáticas en algo que hoy se conoce como Lógica Booleana. Poco después Gottlob Frege, un matemático y filósofo alemán que vivió entre 1848 y 1925, intentó mostrar que la aritmética era parte de la lógica en su libro *Die Grundlagen der Arithmetik* (Fundamentos de la Aritmética). Sin embargo, Georg Cantor, un matemático alemán que vivió entre 1845 y 1918, dio un paso importantísimo en la historia de las matemáticas y de la lógica, incluso adelantándose a Frege en su fundamentación lógica de la aritmética, y creó entre 1878 y 1897 una nueva disciplina matemática: *la teoría de conjuntos.*



Su obra fue admirada y condenada simultáneamente por sus contemporáneos.

Desde entonces los debates en el seno de la teoría de conjuntos, han sido siempre apasionados, sin duda por hallarse estrechamente conectados con importantes cuestiones lógicas.

Según la definición de conjunto de Cantor, éste es "*una colección en un todo de determinados y distintos objetos de nuestra percepción o nuestro pensamiento, llamados los elementos del conjunto*".

Es indiscutible el hecho de que la teoría de conjuntos es una parte de las matemáticas y que es, además, la teoría matemática dónde fundamentar la aritmética y el resto de teorías matemáticas. Es también indiscutible que es una parte de la lógica.

Con la introducción de la teoría de los conjuntos, la enseñanza de las matemáticas cambió sustancialmente. Sin embargo, si bien para el estudiante de ciencias matemáticas, la teoría de los conjuntos resulta apasiónate, no estamos tan seguros acerca de si lo mismo sucede con los estudiantes del colegio secundario.

En la actualidad, muchos tratados de ciencias probabilísticas están basados en la teoría de los conjuntos y para los estudiantes de ciencias sociales, resulta en extremo difícil de entender.

Tal estudiante necesita, entonces, comprender la probabilidad mediante un criterio de lógica, si entendemos por lógica a la ciencia que estudia las formas y las

leyes generales que rigen el pensamiento humano.
Actuar con lógica es una forma de correspondencia con lo razonable, es decir aplicar nuestra capacidad de razonar para actuar con el sentido común que toda persona tiene.
Dentro de este contexto, aparece la lógica matemática, que es una parte de la lógica que emplea en sus operaciones los métodos y el simbolismo de las matemáticas.
Como dijimos, la Lógica estudia la forma del razonamiento y, en este contexto, la Lógica Matemática es la disciplina que trata los métodos de razonamiento. En un nivel elemental, la Lógica proporciona reglas y técnicas para determinar si es o no valido un argumento dado. El razonamiento lógico se emplea en Matemáticas para demostrar teoremas, sin embargo, se usa en forma constante para realizar cualquier actividad en la vida.
En este sentido queremos que el estudiante de ciencias sociales y del comportamiento, entienda la teoría de las probabilidades en la medida y extensión con que la ha de tener que utilizar en su profesión.
Por lo tanto, en este estudio de la probabilidad, no utilizaremos de las matemáticas más que las operaciones básicas y el resto será solo razonamiento y sentido común.

CONCEPTO DE PROBABILIDAD

Desde el sentido común, la probabilidad es la condición de probable y la podemos asociar con todo aquello que no es totalmente seguro de suceder y que, por ende, está gobernado por la incertidumbre. Una incertidumbre que está ligada a nuestro desconocimiento acerca de lo que ha de ocurrir en el futuro.
Desde un punto de vista más matemático, la probabilidad de un suceso[13], es igual al cociente entre el número de casos que pueden ser favorables y el número total de casos posibles (favorables + desfavorables). Como siempre la cantidad de casos favorables es menor a la cantidad de casos posibles, la probabilidad siempre es menor que uno y mayor que cero:

$$0 < p < 1 \qquad \text{[Fórmula 2.1]}^{14}$$

Probabilidad y certeza

La incertidumbre en la ocurrencia de algo, se contrapone con la certeza, es decir

[13] Suceso: Cada uno de los resultados de un fenómeno o experiencia aleatoria. Es sinónimo de evento.
[14] En este libro utilizaremos las letras p y P, en forma indiferente, para indicar la probabilidad

la seguridad en la ocurrencia de ese algo.
Dijimos que la probabilidad estaba dada por un número mayor que cero y menor que uno. Por el contrario, la certeza, tiene solo dos valores: 0 y 1.
Cero, es la certeza acerca de la imposibilidad total de que algo ocurra y uno es la certeza o seguridad total en la ocurrencia del suceso. Por ejemplo:

- la probabilidad de que al arrojar una moneda salga cara o ceca es 1/2 ó 0,5 (50%). Se trata de un valor mayor que cero y menor que uno.
- La probabilidad de que al arrojar una moneda, salga cara o seca, es igual a 1. Es una certeza total pues cara y seca son las dos únicas opciones de caer que tiene la moneda.
- La probabilidad de que al arrojar una moneda al aire, esta quede volando, es igual a cero. Debido a la ley de gravedad, es totalmente imposible que una moneda flote en el aire.

Entonces, un valor de uno, nos indica la certeza total de ocurrencia de un suceso y un valor de cero, nos indica la certeza total del que tal suceso no ha de ocurrir.
En la tabla siguiente se muestran algunas certezas.

Ejemplo	**Certeza**
Después del verano llega el otoño	1
Vivir 300 años	0
Las montañas están sobre el nivel del mar	1
Mantenerse con vida bajo el agua, sin respiración artificial	0
Durante un terremoto, las capas de tierra vibran	1
Que un automóvil naftero funcione sin combustible	0
A temperaturas mayores de 100 grados Celsius , el agua se evapora	1

SUCESOS DETERMINISTAS

Desde el punto de vista de la ocurrencia de los fenómenos de la ciencia, el determinismo es una posición filosófica que sostiene que dicha ocurrencia no es casual ni se produce al azar, sino que obedece a leyes naturales y causales y es debida a la actuación de factores específicos.
Los sucesos o experimentos deterministas, son los experimentos o sucesos de los que podemos predecir el resultado antes de que se realicen.

Ejemplos:

- Si hoy es lunes, mañana será martes.

- Si en este momento es de día, más tarde será de noche.
- Si a un avión se le rompe el sistema de propulsión, caerá.
- Cuando a un líquido se le suministra constantemente calor, entrará en ebullición.
- Si arrojo un dado de seis caras, saldrá un número entre 1 y 6.
- Si juego a la ruleta, saldrá un número entre 0 y 36.

También, podemos decir que un experimento o fenómeno es determinista, si se obtiene el mismo resultado cada vez que se repite el experimento en las mismas condiciones.

SUCESOS ALEATORIOS O ESTOCÁSTICOS[15]

Son aquellos en los que no se puede predecir su resultado, ya que éste depende solo del azar.
Ejemplos:

- Si lanzamos una moneda al aire, no sabemos de antemano si saldrá cara o ceca.
- Si lanzamos un dado, tampoco podemos determinar el resultado que vamos a obtener.
- Si abrimos un libro, sin mirarlo, no sabemos en qué página se abrirá.
- Si vamos al hipódromo desconocemos qué caballo ganará en una carrera honesta.

Otra definición, expresa que un **experimento aleatorio** o **estocástico** es el que puede producir resultados diferentes en las mismas condiciones.

CÁLCULO DE PROBABILIDADES. LÓGICA Y EMPIRISMO



Continúa...

[15] Estocástico, del latín *stochasticus*, que a su vez procede del griego στοχαστικός, "hábil en conjeturar". Fuente: http://es.wikipedia.org/wiki/Estoc%C3%A1stico

Mientras que en teoría, la aleatoriedad es una propiedad intrínseca, en la práctica, es información incompleta.
Nassim Taleb[16]

UNIDAD 3

VARIABLES ALEATORIAS[17]

DETERMINISMO *vs.* ALEATORIEDAD

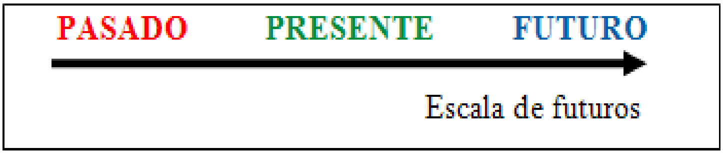
El determinismo es una antiquísima teoría filosófica que expresa que todo lo que ocurre en el universo – incluyendo los pensamientos, emociones y acciones del hombre – ha sido determinado por factores anteriores, de modo que nada podría haber ocurrido de forma diferente a como, de hecho, ocurrió y, además, que todo en el futuro ha sido predeterminado y es inevitable. Todos los aspectos de la vida y el carácter del hombre, según este punto de vista, son simplemente el producto de factores que, en última instancia, están fuera de su control.

---

[16] Nassim Nicholas Taleb. (Nacido en 1960 en el Líbano) es un ensayista, investigador y financiero estadounidense. Es también miembro del Instituto de Ciencias Matemáticas de la Universidad de Nueva York. Se considera a sí mismo «empirista escéptico» y cree que los científicos y los financieros sobreestiman el valor de la explicaciones racionales sobre datos del pasado e infravaloran el peso de la aleatoriedad en esos datos. Fuente: http://es.wikipedia.org/wiki/Nassim_Taleb

[17] Fuente de la imagen: http://e-ducativa.catedu.es/44700165/aula/archivos/repositorio//2001/2009/html/index.html

Esta teoría sostiene que todo acontecimiento está enmarcado en una cadena causal por lo que todo lo que sucede está fuertemente ligado a una causa y en este contexto, el estado actual determina el futuro ya que el presente es el futuro del pasado pero el pasado del futuro.



De alguna manera, como en toda teoría filosófica, existen diferentes niveles de determinismo que se diferencian en los detalles de sus afirmaciones. Las más usuales son:

- El **determinismo fuerte,** que afirma que no existen sucesos genuinamente aleatorios o azarosos y que el futuro es potencialmente predecible a partir del presente, y

- El **determinismo débil,** que afirma que la probabilidad está determinada por los hechos presentes y que existe una fuerte correlación entre los estados presentes y futuros, aun admitiendo la influencia de sucesos esencialmente aleatorios e impredecibles.

Aquí, es importante poner de relieve que existe una diferencia importante entre la determinación y la predictibilidad de los hechos. La determinación implica exclusivamente la ausencia de la aleatoriedad en la cadena causa-efecto responsable de cualquier evento real y concreto
Por su parte, la predictibilidad es un hecho potencial derivado de la determinación certera de los sucesos, pero exige que se conozcan las condiciones iniciales, o de cualquier punto de la cadena de causalidad.

TIPOS DE FENÓMENOS CIENTÍFICOS

La ciencia, en general, estudia dos tipos de fenómenos:

- los determinísticos, y
- los aleatorios o al azar.

Fenómenos determinísticos

Son aquellos en los que podemos predecir su resultado con absoluta precisión, aún antes de ser realizados. Por ejemplo:

- Si lanzamos una piedra al aire, sabemos que pasado un tiempo caerá.
- Si calentamos una barra de metal, sabemos que se dilatará.
- Si enfriamos el agua, sabemos que en un momento se congelará.
- Si bebemos en exceso, sabemos que nos embriagaremos.

Fenómenos aleatorios
Son aquellos en los que no podemos determinar su resultado con exactitud, aunque los repitamos miles o millones de veces. Son ejemplos de fenómenos aleatorios:

- La posición en que caerá una moneda (cara o ceca) lanzada al aires.
- El tiempo que viviremos.
- El número que será ganador en la quiniela de hoy.
- El sexo de nuestro próximo hijo, sin hacer una ecografía.

Las leyes científicas se ocupan de establecer lo que debemos esperar que ocurra en los fenómenos llamados deterministas. Si dejamos caer un dado desde una cierta altura, digamos 20 metros, las leyes de la física nos permitirán determinar cuánto tiempo demorará en llegar al piso y con qué velocidad lo hará. Se trata sin lugar a dudas, de un evento determinista.
Sin embargo, lo que no se podrá conocer con certeza, es cuál es el número que mostrará la cara superior del dado, una vez que se detenga.
Los fenómenos denominados aleatorios tienen como particularidad, la imposibilidad de anticipar con exactitud el resultado del experimento. Sus principales particularidades son:

- Las observaciones se podrán repetir indefinidamente bajo condiciones esencialmente invariables.
- Es posible describir todos los posibles resultados de una observación, pero no es posible establecer cuál será el resultado que ha de ocurrir.
- Los resultados individuales de las observaciones repetidas pueden ocurrir de cualquier forma posible, pero cuando el número de observaciones es grande aparece el patrón de regularidad estadística.

MODELIZACIÓN DE LA REALIDAD
Como venimos siempre diciendo, uno de los objetivos de la estadística es el conocimiento cuantitativo de una determinada parte de la realidad. Para ello, es ne-

cesario construir modelos de esa realidad particular, sabiendo y aceptando que la realidad, o lo real, es siempre más complejo, multifacético y multitemporal que cualquier modelo que se pueda construir. Pero, como la realidad es muy difícil de medir, los modelos constituyen una muy buena aproximación hacia ella dado que la formulación de modelos tiene la gran ventaja de ofrecernos tendencias hacia la realidad que, generalmente, incluye funciones de probabilidad.

No olvidemos que, desde hace muchísimo tiempo, una de las preocupaciones de los científicos ha sido construir modelos de distribuciones de probabilidad que pudieran representar el comportamiento teórico de diferentes fenómenos aleatorios que aparecían en el mundo real. La pretensión de modelar lo observable ha constituido siempre una necesidad básica para el científico empírico, dado que, a través de esas construcciones teóricas, podía experimentar sobre aquello que la realidad no le permitía.

Por otra parte, un modelo resulta extremadamente útil, siempre que se corresponda con la realidad que pretende representar o predecir, de manera que ponga de relieve las propiedades más importantes del mundo que nos rodea, aunque sea a costa de la simplificación que implica todo modelo. Toda distribución de probabilidad es generada por una variable aleatoria X, la que puede ser discreta o continua.

VARIABLE ALEATORIA

Recordemos que una variable aleatoria es aquella en que sus elementos de variación son sucesos de un fenómeno aleatorio, regidos por el azar y que, por ende, está sujeta a una determinada probabilidad de ocurrencia.

Estas variables se dividen en discretas y continuas.

Variable aleatoria discreta

Una variable aleatoria es discreta, cuando sólo puede tomar un número finito de valores dentro de un intervalo. Sus principales características son:

DENOMINACIÓN	CARACTERÍSTICA:	DEBIDO A QUE:
Variable aleatoria discreta (X)	Variable	X, puede tomar diferentes valores.
	Aleatoria	El valor tomado por X, es totalmente al azar.
	Discreta	Los valores que puede adoptar X, son una cantidad finita de valores enteros

Ejemplos:

X =	Valores posibles de *X*
Cantidad de cachorros dados a luz por una perra labradora, en un parto.	1, 2, 3, 4,...n cachorros por parto.
Cantidad de robos anuales en una sucursal bancaria	0, 1, 2, 3,... *n* robos al año
Cantidad de días al año que superan los 20º de temperatura	0, 1, 2, 3,... *n* días al año

Como vemos, los valores que adopta una variable aleatoria discreta no son medibles, son simplemente atributos que solo se pueden contar.

Es importante entender a que nos referimos cuando lo hacemos con respecto a la aleatoriedad. Cuando decimos que los valores de una variable dependen del azar nos referimos, por ejemplo, a que si se lanzan dos dados honesto al aire y *X* representa la cantidad de veces que sale un 6, entonces *X* es una variable aleatoria que puede tomar, al azar, solo los valores 0, 1 ó 2. No existe otra probabilidad, como nos muestra la tabla 3.1.

Una buena definición de un experimento aleatorio es que un experimento es aleatorio si hay más de un resultado posible y no podemos predecir con anterioridad lo que va a suceder. En este caso se dice que el resultado depende del azar.
Podremos concluir, entonces, en que un experimento que tiene las siguientes características es llamado experimento aleatorio o estadístico.

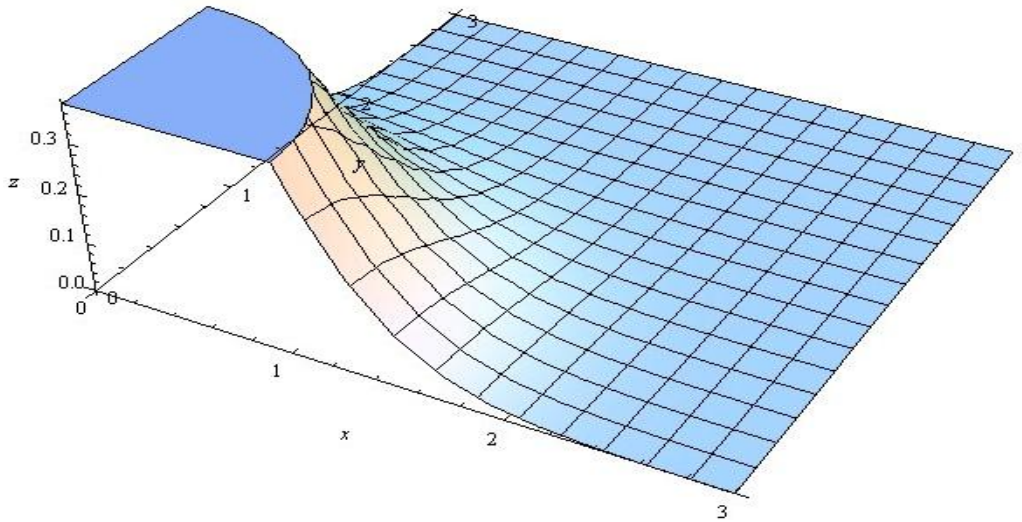
a) Todos los posibles resultados del experimento son conocidos antes de hacer una realización del experimento.
b) El resultado exacto en cualquier ejecución del experimento no es predecible (aleatoriedad)
c) El experimento, teóricamente, puede ser repetido bajo idénticas condiciones.

Continúa ...

Es una verdad muy cierta que, cuando no esté a nuestro alcance determinar lo que es verdad, deberemos seguir lo que es más probable.
Descartes[18], en su Discurso del Método[19]

UNIDAD 4

DISTRIBUCIÓN DE PROBABILIDADES



Los valores de una variable sirven para describir o clasificar individuos o distinguir entre ellos.
Por ejemplo si digo que X es la variable altura y el valor de X para Ernesto es 1,85 metros, sin duda nos imaginamos a Ernesto como un muchacho alto. En-

[18] René Descartes (La Haye, Turena francesa,31 de marzo de 1596 - Estocolmo, Suecia, 11 de febrero de 1650), también llamado Renatus Cartesius, fue un filósofo, matemático y físico francés, considerado como el padre de la geometría analítica y de la filosofía moderna, así como uno de los nombres más destacados de la revolución científica. Fuente: http://es.wikipedia.org/wiki/Ren%C3%A9_Descartes

[19] El Discurso del método (Discours de la méthode en francés), cuyo título completo es Discurso del método para conducir bien la propia razón y buscar la verdad en las ciencias (Discour de la méthode pour bien conduire sa raison, et chercher la vérité dans les sciences) es la principal obra escrita por René Descartes y una obra fundamental de la filosofía occidental con implicaciones para el desarrollo de la filosofía y de la ciencia.
Se publicó de forma anónima en Leiden (Holanda) en el año 1637. Constituía, en realidad, el prólogo a tres ensayos: *Dióptrica*, *Meteoros* y *Geometría*; agrupados bajo el título conjunto de *Ensayos filosóficos*. Fuente: http://es.wikipedia.org/wiki/Discurso_del_m%C3%A9todo

tonces, la variable altura, nos sirve para describir a los integrantes de una muestra. con respecto a sus alturas. La mayoría de nosotros hacemos algo más que simplemente describir, clasificar o distinguir, porque tenemos ideas respecto a las *frecuencias relativas* de los valores de una variable. En estadística decimos que la variable tiene una *función* de *probabilidad*, una *función* de *densidad* de *probabilidad* o simplemente una *función de distribución.*

Las distribuciones de probabilidad están relacionadas con la distribución de frecuencias y, por tal motivo, podemos pensar en la distribución de probabilidad como una distribución de frecuencias teórica. Una distribución de frecuencias teórica es una distribución de probabilidades que describe la forma en que se espera que varíen los resultados. Debido a que estas distribuciones tratan sobre expectativas de que algo sucederá, resultan ser modelos útiles para hacer inferencias y tomar decisiones de incertidumbre.

Podemos decir, entonces, que los objetivos de las distribuciones de probabilidad son:

a) Introducir las distribuciones de probabilidad que más se utilizan en la toma de decisiones.
b) Mostrar qué distribución de probabilidad podemos utilizar, y cómo encontrar sus valores.
c) Entender las limitaciones de cada una de las distribuciones de probabilidad que utilicemos.

DISTRIBUCIÓN DE PROBABILIDADES

Una distribución de probabilidades indica toda la gama de valores que pueden ocurrir como resultado de un experimento, en el caso que éste se llevase a cabo. Es decir, describe la probabilidad de que un evento pueda suceder en el futuro y, por tal motivo, constituye una herramienta fundamental para la prospectiva, dado que se puede diseñar un escenario de acontecimientos futuros considerando las tendencias actuales de diversos fenómenos naturales.

Las variables descriptas anteriormente nos generan distribuciones de probabilidad, las que pueden ser:

- Discretas, o
- Continuas

a) **Distribuciones de probabilidad discretas**

Las características de esta distribución, son:

a.1. Es generada por una variable aleatoria discreta (X).

a.2. Las probabilidades asociadas a cada uno de los valores que toma X deben ser mayores o iguales a cero. $\Rightarrow P(X_i) \geq 0$

a.3. La sumatoria de las probabilidades asociadas a cada uno de los valores que toma X debe ser igual a 1. $\Rightarrow \Sigma P(X_i) = 1$

b) **Distribuciones de probabilidad continuas**

Las características de esta distribución, son:

b.1. Es generada por una variable aleatoria continua (X).

b.2. Las probabilidades asociadas a cada uno de los valores que toma X deben ser mayores o iguales a cero. $\Rightarrow P(X_i) \geq 0$.
Dicho de otra forma, la función de densidad de probabilidad deberá tomar solo valores mayores o iguales a cero.

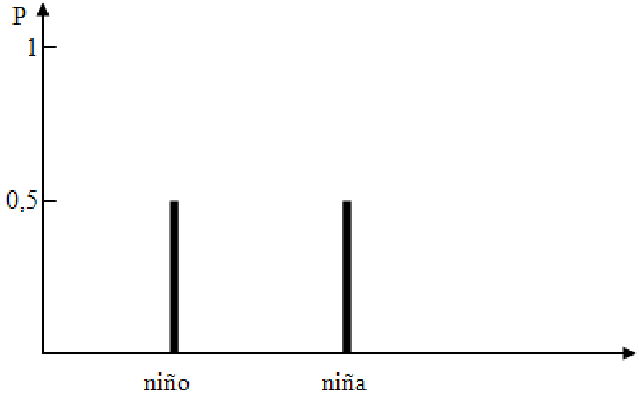
b.3. La sumatoria de las probabilidades asociadas a cada uno de los valores que puede adoptar X debe ser igual a 1. $\Rightarrow \Sigma P(X_i) = 1$.

Dicho de otra manera, el área definida bajo la función de densidad de probabilidad deberá ser unitaria (igual a 1).
Hasta ahora todos los ejemplos que hemos visto en las unidades anteriores fueron para calcular la probabilidad de ocurrencia de un determinado evento.
Veremos, ahora, como aplicamos el cálculo de probabilidades a lo que se conoce como distribuciones de probabilidad, sean estas continuas o discretas. La distribución de probabilidad de una variable aleatoria, es una función que asigna a cada suceso definido sobre la variable aleatoria, la probabilidad de que dicho suceso ocurra o no. Esta distribución de probabilidad está definida sobre el conjunto de todos los sucesos, donde cada uno de los sucesos es el rango de valores de la variable aleatoria y está completamente especificada por la función de distribución, cuyo valor para cada valor de la variable aleatoria X, es la probabilidad de que la variable aleatoria sea menor o igual que X.

GRAFICACIÓN DE LAS DISTRIBUCIONES DE PROBABILIDAD

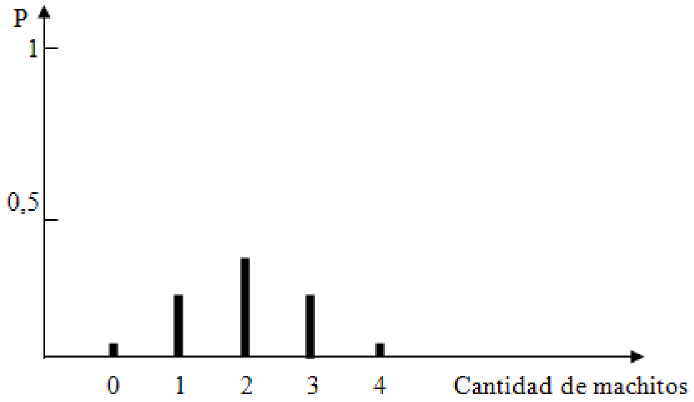
Si representamos gráficamente la probabilidad de nacimiento de un niño o de una niña en un mismo parto, o la salida de cara o ceca en el lanzamiento de una única moneda o de la salida de un tres o un cinco en el lanzamiento de un único dado, siempre la probabilidad de éxito y de fracaso, como se denomina en estadística a los casos contrarios, es la misma. Se trata de lo que se conoce como sucesos equiprobables y su graficación sería del siguiente tipo:

Gráfico 4.1. Probabilidad de nacimiento de un niño o una niña



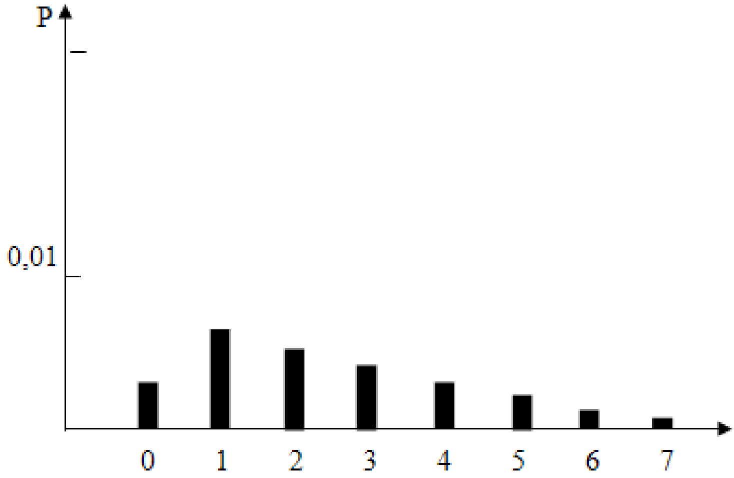
Supongamos ahora que queremos graficar la probabilidad de nacimiento de: ningún machito o un machito o dos machitos o tres machitos o cuatro machitos, en un parto múltiple (cuatro gatitos) de una gatita siamesa. El gráfico 4.2, nos muestra como sería tal distribución de probabilidades

Gráfico 4.2. Probabilidad de nacimiento de 0, 1, 2, 3 ó 4 gatos machitos en un parto único de cuatro gatitos.



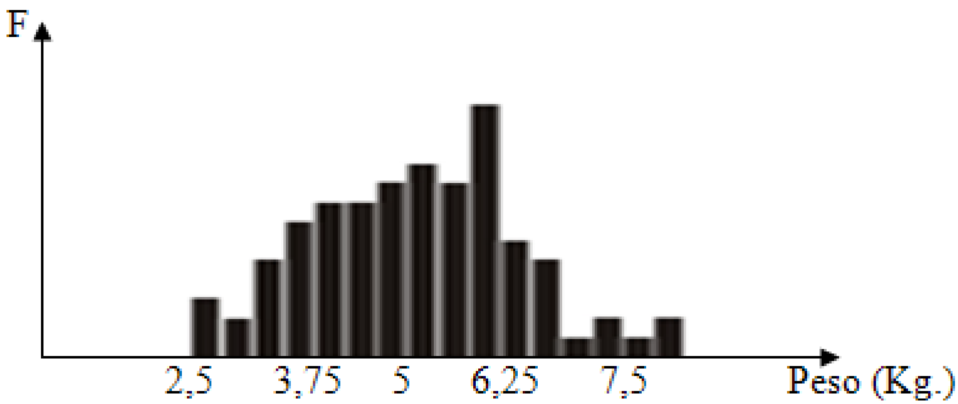
Sabemos que la probabilidad de que nazca un gato color venus, es mus baja. Si quisiéramos graficar cual es la probabilidad de que en un hospital veterinario, donde nacen 20.000 gatos por año, un día determinado nazcan entre 0 y 7 gatitos color venus, obtendríamos un gráfico como el 4.3.

Gráfico 4.3. Probabilidad de que en un día determinado
nazcan entre 0 y 7 gatitos color venus



Y si quisiéramos graficar el peso de los mil gatos adultos que se encuentran internados en de un determinado hospital veterinario, el gráfico sería similar al histograma que se muestra en la figura 4.4

Gráfico 4.4. Peso de mil gatos adultos



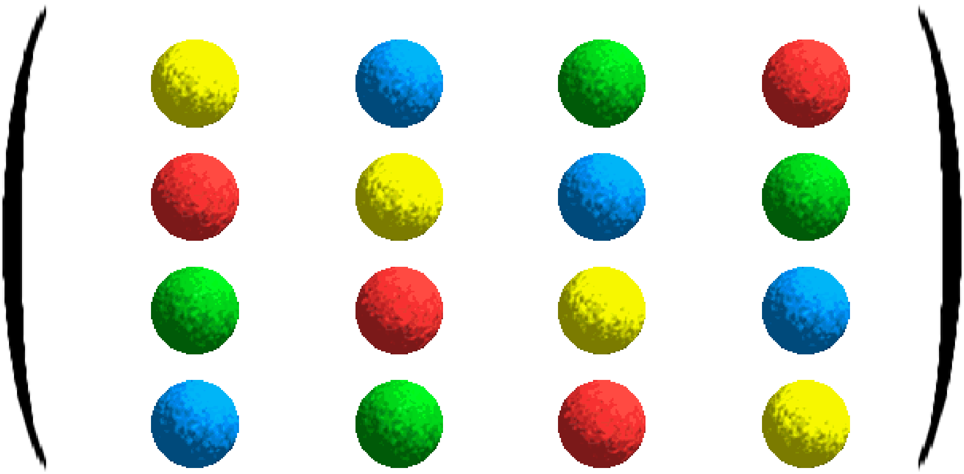
Pero, si (hipotéticamente) quisiéramos representar el peso de todos los gatos de una ciudad, tendríamos un gráfico con forma de campana como el que nos

Continúa ...

La Matemática tiene la virtud de elevar el alma,
obligándola a razonar acerca de los números.
Platón[20]

UNIDAD 5

NOCIONES DE ÁLGEBRA LINEAL
VECTORES Y MATRICES[21]



Durante el desarrollo del libro, hemos tratado de explicar y aplicar los conceptos y cálculos de probabilidad mediante un uso amigable y mínimo de las matemáticas. Esta unidad tiene por finalidad estudiar, también, los conceptos mínimos y necesarios de álgebra lineal para poder comprender, en la unidad siguiente, el proceso aleatorio de Márkov.
El álgebra lineal es una parte del cálculo matemático que se dirige, principalmente, al desarrollo y utilización de ecuaciones y sistemas de ecuaciones lineales, vectores y matrices. Los conceptos y métodos del álgebra lineal han con-

[20] Platón: Filósofo griego (Atenas, 427 - 347 a. C.). Nacido en el seno de una familia aristocrática, abandonó su vocación política por la Filosofía, atraído por Sócrates. Siguió a éste durante veinte años y se enfrentó abiertamente a los sofistas (Protágoras, Gorgias…). Tras la muerte de Sócrates (399 a. C.), se apartó completamente de la política; no obstante, los temas políticos ocuparon siempre un lugar central en su pensamiento, y llegó a concebir un modelo ideal de Estado. Viajó por Oriente y el sur de Italia, donde entró en contacto con los discípulos de Pitágoras; luego pasó algún tiempo prisionero de unos piratas, hasta que fue rescatado y pudo regresar a Atenas. Fuente: http://www.biografiasyvidas.com/biografia/p/platon.htm
[21] Fuente de la imagen: http://thales.cica.es/rd/Recursos/rd99/ed99-0289-02/ed99-0289-02.html

buido decisivamente al desarrollo de muchas otras áreas del conocimiento tanto dentro como fuera de la Matemática. Entre muchas de ellas podemos mencionar la teoría cualitativa de ecuaciones diferenciales, la teoría de códigos y ía, la teoría de decisiones, la robótica, la astronomía y la programación lineal. No es exagerado afirmar que sus ideas y resultados aparecen en casi todo el llo humano.

Por eso, estimados lectores, a no asustarse al leer esta unidad. Aprendamos a no odiar los números, pues, como dice Platón, ellos nos enaltecerán el alma, lo cual, no es poca cosa.

ORÍGENES DEL ÁLGEBRA LINEAL

"Los primeros rudimentos de lo que hoy conocemos como Álgebra lineal se han

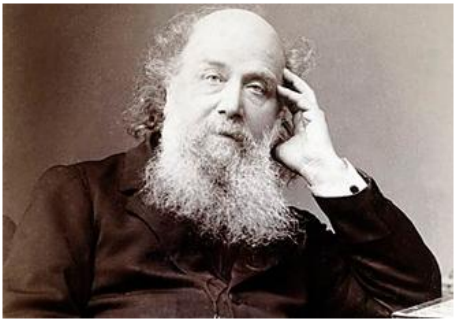
encontrado en el documento matemático más antiguo que ha llegado hasta nuestros días: el papiro Rhind[22], conservado en el British Museum con algunos fragmentos en el Brooklyn Museum, y conocido también como el Libro de Cálculo, el cual fue escrito por el sacerdote egipcio Ahmés hacia el año 1650 a.C. y exhumado en Tebas en 1855 ([11], Vol. I, pág. 40). En este valioso documento se consideran las ecuaciones de primer grado, donde la incógnita aparece representada por un "ibis" que significa escarbando en el suelo, posiblemente por su primogénita aplicación a la agrimensura. Este documento contiene 85 problemas redactados en escritura hierática y fue concebido originalmente como un manual práctico para los no iniciados. Según el propio Ahmés, este texto es una copia de uno más antiguo (2000-1800 a.C.), algunos de cuyos documentos proceden quizá de períodos más antiguos.

Los babilonios sabían cómo resolver problemas concretos que involucraban ecuaciones de primer y segundo grado, usando completación de cuadrados o sustitución, así como también ecuaciones cúbicas y bicuadráticas, y sistemas de

22 Fuente de la imagen: http://puemac.matem.unam.mx/puemaco/cuadratura/html/index.html

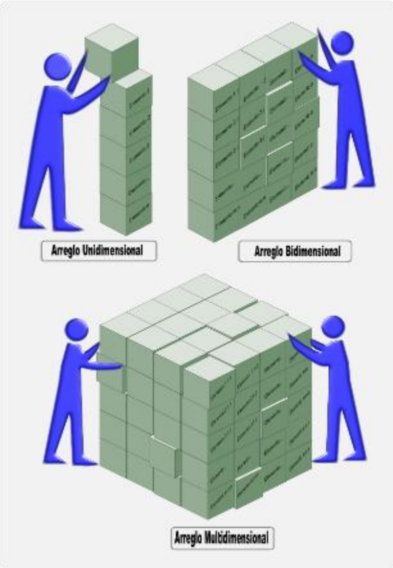
ecuaciones lineales y no lineales."[23]

MATRICES Y VECTORES



Las matrices, no son algo nuevo. El primero en usar el término "matriz" fue el matemático inglés James Joseph Sylvester[24] (1814-1897) en 1850, quien definió una matriz como un "*oblong arrangement of terms"* (arreglo cuadrilongo de términos).[25] Hoy se define de manera muy similar: una matriz es un arreglo[26] bidimensional o tabla bidimensional de números, consistente en cantidades abstractas que pueden sumarse y multiplicarse entre sí. Los números que conforman la matriz son sus entradas, es decir, las entradas de la matriz.

Otra forma de definirla, es como una tabla cuadrada o rectangular de datos (llamados elementos o entradas de la matriz) ordenados en filas y columnas, donde una fila es cada una de las líneas horizontales de la matriz y una columna es cada una de las líneas verticales de la matriz.



A una matriz con *m* filas y *n* columnas se le denomina matriz *m* x *n* donde *m* y *n* son las dimensiones de la matriz y siempre son ≥ 0. Siempre la matriz se denomina en primer término con la cantidad de filas y luego con la cantidad de columnas. Ver la figura 5.1.

El número o la cantidad de elementos de una matriz lo obtendremos de multiplicar el

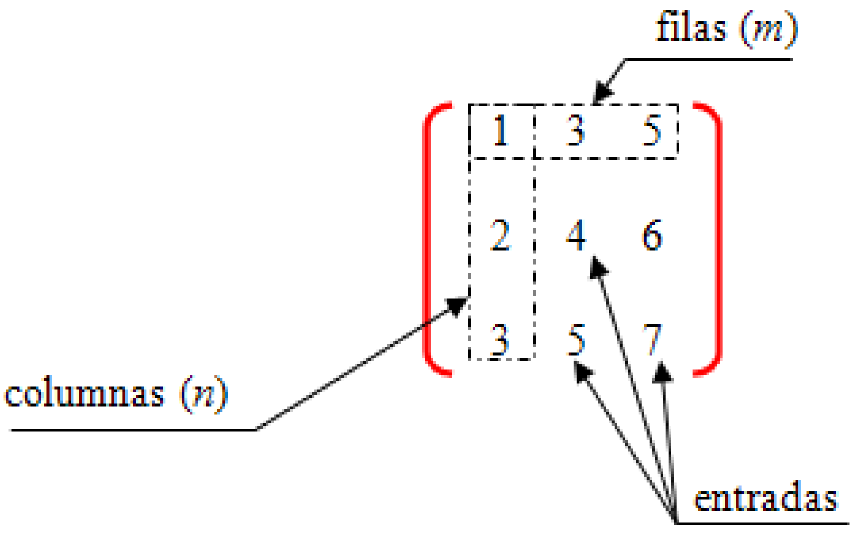
[23] Fuente: http://www.emis.de/journals/DM/v14-2/art6.pdf
[24] Fuente de la imagen: http://cjch.cl/2012/09/semana-3-al-9-de-septiembre-2012/
[25] Fuente: http://www.emis.de/journals/DM/v14-2/art6.pdf
[26] N. del A. Por arreglo se entiende a las estructuras de datos más sencillas. Los arreglos de n objetos tomados de a *k* son las sucesiones de k términos diferentes que se pueden formar con los objetos dados. Por ejemplo, los arreglos de los números 1, 2 y 3, tomados de a tres son: 123, 132, 213,231, 312 y 321. Los arreglos de los mismos números tomados de a dos, son: 12, 13, 21, 23, 31 y 32. Por su parte los arreglos de los mismos números tomados de a uno, son: 1, 2 y 3. Fuente de la imagen:
http://aprendeenlinea.udea.edu.co/lms/ocw/course/view.php?id=9&topic=7

número de filas por el de columnas (m x n). Al producto m x n lo llamamos orden de la matriz.

Orden	**Cantidad de filas**	**Cantidad de columnas**
m x n	m	n
m x 1	m	1
1 x n	1	n
m	m	m

Por lo general se trabaja con matrices formadas por números reales.
Se dice que dos matrices son iguales, cuando tienen el mismo tamaño (igual cantidad de filas y columnas) y las mismas entradas. Cuando hablamos de vectores, nos estamos refiriendo a una matriz de una única fila (1x n) o de una única columna (m x 1). Cuando una matriz tiene igual cantidad de filas (n) y columnas (m), se dice que es una matriz cuadrada (matriz cuadrada de orden n), cuando n ≠ m, las matrices son rectangulares.

Figura 5.1. Características de una matriz (3 x 3)



Si quisiéramos analizar, por ejemplo, los fenómenos del medioambiente que nos circunda, resulta necesario estudiar los datos que describen su comportamiento. La toma y posterior análisis de esos datos representa una parte esencial del traba-

jo que debe realizar el investigador. Los valores tomados en el trabajo de campo contendrán la información indispensable que servirá de base para el posterior análisis.
Extraer esa información no siempre es tarea sencilla y precisamos, entonces, la ayuda de métodos más sofisticados. Las matemáticas y la estadística aportan las herramientas clave que permiten revelar las estructuras internas de los datos que no son accesibles mediante la observación directa.
Estas estructuras básicas a las que nos referimos son las *uplas* o vectores y las matrices. Veamos un ejemplo en el que se introducen los conceptos de *upla* y matriz como herramientas para manipular la información de cierto problema.

NOTACIÓN MATRICIAL

Como dijimos, Una matriz m × n es una tabla de m filas y n columnas de escalares. Es decir, un objeto de la forma:

$$A= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ a_{n1} & a_{n2} & & a_{mn} \end{pmatrix}$$

Donde cada a_{ij} es un escalar. Hasta ahora hemos hablado de números, sin especificar de qué tipo de números se trataba. En esta unidad utilizaremos números racionales ($\mathcal{Q}$), reales ($\mathcal{R}$) o complejos ($\mathcal{C}$). Por lo tanto, a partir de ahora, en vez de números diremos escalares, y al conjunto de números que estemos utilizando lo llamaremos cuerpo de escalares o simplemente cuerpo. Concretando, por ahora basta con pensar que un escalar es un número racional, real o complejo.

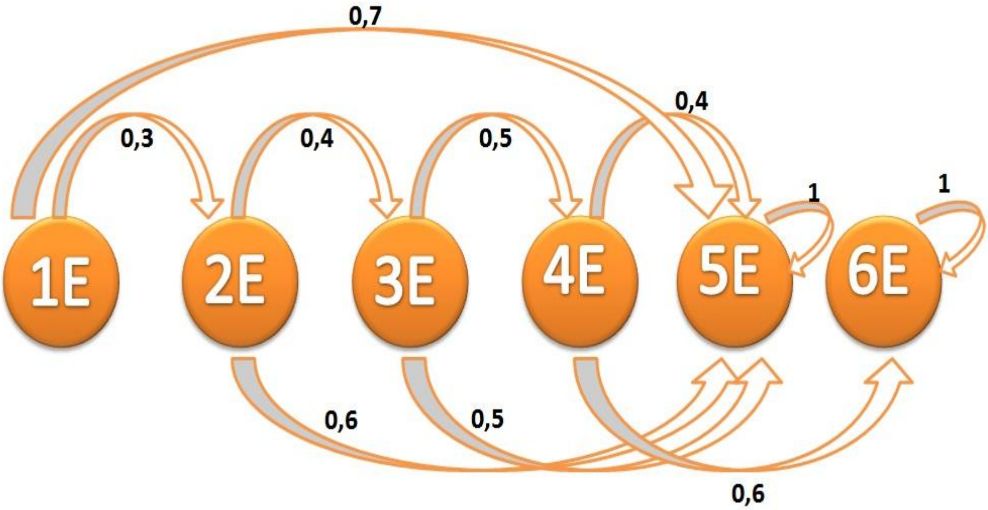
Denotaremos $\boldsymbol{M}_{m\times n}$ (K) al conjunto de matrices $m\times n$, cuyo cuerpo de escalares es K. Si no nos interesa especificar el cuerpo de escalares, escribiremos simplemente $\boldsymbol{M}_{m\times n}$.
Normalmente usaremos una letra mayúscula para denotar una matriz, y la misma letra en minúscula, con los subíndices correspondientes, para denotar sus

Continúa ...

Todos los modelos son erróneos,
pero algunos son útiles.
George E. P. Box[27]

UNIDAD 6

INTRODUCCIÓN A LAS CADENAS DE MÁRKOV FINITAS[28]



En matemáticas, una cadena es, simplemente, una sucesión de elementos.
La denominada cadena de Márkov es un modelo matemático dinámico y estocástico que nos permite describir la *evolución* de un sistema a lo largo del tiempo, a través de la probabilidad.
También, podríamos decir que una cadena de Márkov es una sucesión de ensayos u observaciones análogas en la cual cada ensayo tiene el mismo número finito de resultados posibles y en donde la probabilidad de cada resultado, para un ensayo dado, depende sólo

[27] George Edward Pelham Box (18 octubre 1919 - 28 marzo 2013) fue un estadístico inglés, quien trabajó en las áreas de control de calidad, análisis de series de tiempo, diseño de experimentos e inferencia bayesiana. Fuente: http://en.wikipedia.org/wiki/George_E._P._Box
[28] Fuente de la imagen:
http://investigaciondeoperaciones2ingind.blogspot.com.ar/2011/06/cadenas-de-márkov-continuacion.html

del resultado del ensayo inmediatamente precedente y no de cualquier resultado previo.
Estas cadenas son el resultado de los trabajos de investigación matemática realizados por el matemático de origen ruso Andrei A. Márkov[29] entre los años 1907 y 1912 y que en la actualidad han encontrado muchas aplicaciones en diversos campos del conocimiento.
En la mayoría de los procesos estocásticos, cada resultado depende de lo que sucedió en etapas anteriores del proceso. Por ejemplo:

- El tiempo, en un día determinado, no es aleatorio por completo sino que es afectado, en cierto grado, por el tiempo de los días previos.
- El precio de una acción, al cierre de cualquier día, depende en cierta medida del comportamiento de la bolsa en los días previos.

Más allá de que la utilización de las cadenas de Márkov nos permite encontrar la probabilidad de que un sistema se encuentre en un estado en particular en un momento dado, resulta muy importante que también nos permite encontrar el promedio “a *la larga*” o las probabilidades de estado estable para cada estado, lo cual nos permite predecir el comportamiento del sistema a través del tiempo.
El caso más simple de un proceso estocástico en el que los resultados dependen de otros, ocurre cuando el resultado en cada etapa sólo depende del resultado de la etapa anterior y no de cualquiera de los resultados previos.
Tal proceso se denomina proceso de Márkov o cadena de Márkov, por tratarse de una cadena de eventos donde cada evento está ligado al precedente.
Estas cadenas tienen memoria, ya que recuerdan el último evento y eso condiciona las posibilidades de los eventos futuros. Esto justamente las distingue de una serie de eventos independientes como el hecho de tirar un dado o lanzar una moneda al aire.
Este tipo de proceso presenta una forma de dependencia simple, pero muy útil en muchos modelos, entre las variables aleatorias que forman un proceso estocástico.
Esta herramienta de análisis a pesar de haber surgido a principios del siglo pasado es, actualmente, una técnica novedosa de gran utilidad y rigor científico que sirve tanto para economistas, sociólogos, físicos, y cualesquiera otros investigadores que analicen un proceso dado bajo ciertas circunstancias probabilísticas. La tabla siguiente nos muestra algunos ejemplos de la utilización de las cadenas

[29] La biografía de Andrei Márkov, se muestra en el anexo 5.1.

de Márkov para resolver diversos problemas, como ser[30]:

Área	Utilización
Física	Para resolver problemas de termodinámica y física estadística.
Meteorología	En la formulación de modelos climáticos. Si consideramos el clima de una región a través de distintos días, es claro que el estado actual solo depende del último estado y no de toda la historia en sí, de modo que se pueden usar cadenas de Markov para formular modelos climatológicos básicos.
Epidemiología	En la formulación de modelos matemáticos de epidemias. Una importante aplicación de las cadenas de Markov se encuentra en el proceso Galton-Watson. Éste es un proceso de ramificación que se puede usar, entre otras cosas, para modelar el desarrollo de una epidemia.
Juegos de azar	Son muchos los juegos de azar que se pueden modelar a través de una cadena de Markov. El modelo de la ruina del jugador, que establece la probabilidad de que una persona que apuesta en un juego de azar finalmente termine sin dinero, es una de las aplicaciones de las cadenas de Markov en este rubro.
Finanzas	En la evaluación de opciones para determinar cuándo existe oportunidad de arbitraje, así como en el modelo de colapsos de una bolsa de valores o para determinar la volatilidad de los precios.
Genética	En teoría de genética de poblaciones, se utiliza para describir el cambio de frecuencias génicas en una población pequeña con generaciones discretas, sometida a deriva genética.
Internet[31]	El *pagerank* de una página web (usado por Google en sus motores de búsqueda) se define a través de una cadena de Márkov, donde la posición que tendrá una página en el buscador será determinada por su peso en la distribución estacionaria de la cadena.

[30] Ejemplos tomados de: http://es.wikipedia.org/wiki/cadena_de_M%C3%A1rkov

[31] En http://blog.kleinproject.org/?p=1605&lang=es hay un artículo muy interesante titulado: *CÓMO FUNCIONA GOOGLE: CADENAS DE MÁRKOV Y VALORES PROPIOS.* A quienes se interesen por este tema, se les recomienda su lectura.

Música	Diversos algoritmos de composición musical usan cadenas de Márkov, por ejemplo el software Csound o Max.
Operaciones	Se emplean cadenas de Márkov en inventarios, mantenimiento, flujo de procesos, planes de mantenimiento y reemplazo de equipos, etc.
Genética	Para el estudio de la evolución de poblaciones.
Servicios hospitalarios	Para la toma de decisiones en la administración de salud.
Simulación	Las cadenas de Markov son utilizadas para proveer una solución analítica a ciertos problemas de simulación tales como el Modelo M/M/1.

La aplicación y resolución pormenorizada de las cadenas de Márkov exceden las finalidades de este texto por cuanto es un modelo dinámico y estocástico que se estudia y desarrolla en profundidad en los cursos de matemática pura que muchas veces requieren la utilización de programas informáticos de cálculo matemático como Matlab, Derive y otros.

CONCEPTO DE CADENA DE MÁRKOV FINITA

Una cadena de Márkov para la que existe sólo un número finito k de estados posibles $S_1, \ldots, S_k$ y en cualquier instante de tiempo la cadena está en uno de estos k estados, se denomina cadena de Márkov finita.

PROCESO ESTOCÁSTICO DE LA CADENA DE MÁRKOV

Una cadena de Márkov es un proceso estocástico en el que:

Si el estado actual X_n y los estados previos $X_1, \ldots, X_{n-1}$, ***son conocidos***

Entonces, la probabilidad del estado futuro X_{n+1}

No depende de los estados anteriores $X_1, \ldots, X_{n-1}$

Solamente depende del estado actual X_n

Los valores de las distintas probabilidades p_{ij} son llamadas probabilidades de transición y pueden ordenarse en una matriz, que también se denomina matriz de transición.

$$\boldsymbol{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nm} \end{pmatrix}$$

Esta matriz ***P*** de transición de la cadena de Márkov, es también una matriz estocástica. Veamos algunos ejemplos sencillos para ver como se arma la matriz de transición.

Ejemplo 6.1

Un comerciante va todos los días a trabajar en su auto (A) o en taxi (T). Los estados previos son los siguientes:

- Nunca toma taxi dos días seguidos.
- Si va a trabajar en su automóvil, al día siguiente puede ir nuevamente a trabajar en su automóvil o en taxi en forma indistinta.

El espacio de estados del sistema es [T , A]

Comprobación: El proceso al que hemos hecho referencia es una cadena de Márkov ya que los resultados futuros dependen únicamente de lo que ha pasado el día actual.

La matriz de transición queda de la siguiente manera:

$$\begin{array}{ccc} & & \begin{array}{cc} T & A \end{array} \longrightarrow \text{Hechos futuros} \\ \text{Hechos actuales} \longleftarrow & \begin{array}{c} T \\ A \end{array} & \begin{bmatrix} 0 & 1 \\ 1/2 & 1/2 \end{bmatrix} \end{array}$$

La primera fila de la matriz, corresponde al hecho de que nunca va a trabajar en

Continúa ...