EL CONCEPTO DE PROBABILIDAD

SUCESOS CIERTO, IMPOSIBLE Y ALEATORIO.- Sobre la base de la observación y la experimentación las ciencias enuncian leyes que explican y gobiernan los fenómenos bajo estudio. La más elemental y general de estas leyes podría enunciarse de la siguiente manera:

1. <u>En cada realización de un conjunto de condiciones $\mathcal{G}$ ocurre el suceso A.</u>

En efecto, para ilustrar esta afirmación general supongamos calentamos a 100° C y a la presión atmosférica, 760 mm (conjunto de condiciones $\mathcal{G}$), agua, el resultado será que dicha agua se transformará en vapor (suceso A). Análogamente, para cualquier reacción química de substancias sin intercambio con el medio que le rodea (condiciones $\mathcal{G}$), la cantidad total de materia permanece constante (suceso A). Esta afirmación es conocida como la ley de conservación de la materia. Fenómenos como los descritos son sobradamente conocidos en Química, Física, Biología y otras ciencias experimentales.

Los sucesos que resultan de la realización de determinado conjunto de condiciones pueden agruparse en tres tipos:

a) Aquellos sucesos que inevitablemente ocurren para cada realización del conjunto de condiciones $\mathcal{G}$ y que se denominan sucesos <u>ciertos</u> o <u>seguros</u>

b) Si A no puede ocurrir cuando se realizan las condiciones $\mathcal{G}$, se dice que A es un suceso <u>imposible</u>, finalmente

c) Si realizadas las condiciones $\mathcal{G}$, A puede o no puede ocurrir, se dice entonces que A es un suceso <u>aleatorio</u>.

** Obsérvese que de estas definiciones los conceptos de <u>certeza</u>, <u>imposibilidad</u> y <u>aleatoriedad</u> están siempre referidos a un conjunto definido de condiciones $\mathcal{G}$.

La afirmación de aleatoriedad para un suceso A aporta muy poca información para el conocimiento del suceso, establece simplemente que el conjunto de condiciones $\mathcal{G}$ no refleja la colección entera de razones necesarias y suficientes para la realización de A. En nuestro interés por conocer más acerca de este tipo de sucesos, observamos que existen una gran cantidad de fenómenos en los que, dada una serie de repeticiones de las realizaciones de las condiciones $\mathcal{G}$, la proporción de ocurrencias de A sólo ocasionalmente se desvía de forma significativa de un cierto valor medio, valor que puede servirnos para caracterizar de alguna manera el suceso examinado. Para este tipo de fenómenos podemos, no sólo establecer su aleatoriedad, sino estimar cuantitativamente la posibilidad de su ocurrencia. Esta estimación podemos expresarla mediante una proposición del tipo:

2. <u>La probabilidad de que un suceso A ocurra durante la realización de un conjunto de condiciones $\mathcal{G}$ es igual a</u> p.

Regularidades de este tipo son conocidas como <u>probabilísticas</u> o <u>estocásticas</u> y juegan un importante papel en diversos campos de la ciencia. Por ejemplo, no hay forma de predecir si un determinado átomo de radio se desintegrará o no en un período de tiempo dado, pero es

posible, sobre la base de experimentos de ajuste, determinar la probabilidad de tal desintegración y decir: un átomo de radio se desintegra en un intervalo de tiempo de t años con una probabilidad

$$p = 1 - e^{-0.000436t}$$

El conjunto de condiciones $\mathcal{G}$ consiste, aquí, en el hecho de considerar un átomo de radio, durante un período de t años, no sujeto a ninguna acción externa inusual (como bombardeo con partículas de alta velocidad, por ejemplo), cualesquiera otras condiciones de existencia son irrelevantes en el proceso que nos ocupa (presión, temperatura, etc..). El <u>suceso A consiste en el hecho de que el átomo se desintegre en un período dado de t años</u>.

La idea que parece ahora muy natural, de que la probabilidad de un suceso aleatorio A, bajo condiciones conocidas, admite una cuantificación mediante algún número $p = P(A)$ fue elaborada de manera sistemática por primera vez en el siglo 17 en los trabajos de Fermat (1601-1665), Pascal (1623-1662), Huygens (1629-1695) y en particular por Jacobo Bernoulli (1654-1705). Sus investigaciones colocaron los fundamentos de la teoría de la probabilidad. Desde aquel tiempo, esta teoría ha estado desarrollándose como una disciplina matemática y ha sido enriquecida con importantes resultados. Su aplicación al estudio de fenómenos reales de la más diversa naturaleza continúa encontrando nuevas e importantes confirmaciones.

No existe ninguna duda en cuanto a que el concepto de probabilidad matemática comporta connotaciones filosóficas que justifican su estudio desde esta perspectiva. El problema básico, desde la perspectiva filosófica, sería el siguiente: <u>bajo qué condiciones existe significación objetiva en la estimación cuantitativa de la probabilidad de un suceso aleatorio A con la ayuda de un número fijo P(A), llamado probabilidad matemática del suceso A, y cuál es la significación objetiva de esta estimación</u>?

Cada investigador que aplica la teoría de la probabilidad a su área específica, procede realmente con la convicción de que <u>los juicios probabilísticos expresan ciertas propiedades objetivas del fenómeno bajo estudio</u>. Afirmar que bajo ciertas condiciones $\mathcal{G}$ la ocurrencia del suceso A tiene una probabilidad p es afirmar que entre el conjunto de condiciones y el suceso existe una cierta relación perfectamente definida, aunque peculiar y no por ello menos objetiva, relación que es independiente del investigador. El problema es determinar la naturaleza de esta relación. La dificultad de contestar a este último interrogante ha hecho posible la circunstancia paradójica de que algunos científicos hayan afirmado que los juicios probabilísticos dependen sólo del estado del investigador. Hay que señalar, que si bien esto no es así, la experiencia demuestra que el significado objetivo de la probabilidad matemática tiene sentido en el contexto de un conjunto de condiciones perfectamente definido.

En cualquier caso cuestiones de este tipo no deben ser objeto de estudio en un curso como el presente y nos ocuparemos tan sólo de los aspectos matemáticos del problema al margen de la determinación de la naturaleza del concepto.

DIFERENTES APROXIMACIONES A LA DEFINICIÓN DE PROBABILIDAD

Un gran número de definiciones de la probabilidad matemática han sido propuestas por diferentes autores. No nos ocuparemos aquí de todas ellas. Hay que señalar que el interés por una definición lógicamente irreprochable del concepto de probabilidad surgió más tarde, históricamente hablando, que la capacidad para determinar la probabilidad de los sucesos, para llevar a cabo cálculos con estas probabilidades y también para utilizar los resultados de estos cálculos en asuntos prácticos y en investigación científica. Por esta razón, en gran parte de los intentos por establecer una definición científica del concepto general de probabilidad es fácil percibir varios aspectos del proceso de conocimiento concreto que en cada caso específico conduce a una determinación real de la probabilidad de un suceso dado, ya se trate de la probabilidad de obtener un seis en cuatro lanzamientos de un dado o la probabilidad de desintegración de un átomo de una materia radiactiva.

En la gran mayoría de las definiciones podemos establecer los tres grupos siguientes:

1.- Definiciones de la probabilidad matemática como una medida cuantitativa del <u>grado de certidumbre</u> del observador.

2.- Definiciones que reducen el concepto al de <u>igual posibilidad</u>, el más primitivo de todos (se las conoce como <u>definiciones clásicas</u>).

3.- Definiciones que proceden a partir de la <u>frecuencia</u> de ocurrencia de un suceso en un elevado número de pruebas (<u>definiciones estadísticas</u>).

Nos ocuparemos de las del tipo 2 y 3 no sin antes criticar las definiciones agrupadas en el tipo 1. En efecto, el hecho de admitir la existencia de la realidad al margen de nosotros y del conocimiento que de ella tengamos y de tener en cuenta el éxito de los juicios probabilísticos en el conocimiento y aprendizaje del mundo exterior, debería dejar suficientemente claro que las definiciones puramente subjetivas de la probabilidad matemática son totalmente insostenibles. El problema es que en el lenguaje ordinario las expresiones "probablemente", "muy probablemente", "es altamente improbable", etc. expresan simplemente la actitud del que habla respecto de la verdad o la falsedad de algún juicio <u>simple</u>, en contra de otras interpretaciones que justificarían definiciones subjetivas como las del tipo 1. Debemos por tanto poner el énfasis en una circunstancia que todavía no ha sido señalada: <u>mientras entodas las aplicaciones científicas</u> de la teoría de la probabilidad "probabilidad" es la probabilidad de la ocurrencia de algún suceso A, cuando un determinado conjunto de condiciones $\mathcal{G}$, que es <u>fundamentalmente reproducible un infinito número de veces</u>, ha sido realizado (es solo en tal situación que la afirmación $p = P(A)$ expresa una cierta regularidad con significado objetivo), <u>en el habla ordinaria</u> es costumbre hablar de una mayor o menor probabilidad de algún juicio <u>muy definido</u>. Veamos un ejemplo a través de los dos juicios siguientes:

a) cualquier número natural, par, mayor que 2 puede ser representado mediante la suma de dos números primos ($4=2+2$, $6=3+3$, $8=5+3$, ...)

b) nevará en Valencia el 3 de Febrero de 1982.

Respecto de la afirmación hecha en a) si bien no tenemos un conocimiento completo muchos se inclinarían a creerlo como extremadamente probable, en cuanto a la afirmación hecha en b) habrá que esperar al 3 de Febrero de 1982 para tener una respuesta exacta. Sin embargo y dado que no suele nevar en Valencia nunca, esta afirmación será considerada altamente improbable. El concepto de probabilidad utilizado en el análisis de estos dos juicios no tiene nada que ver con el correspondiente concepto matemático. En ambos casos se trata de reflejar la actitud del que habla respecto de la cuestión planteada, actitud que tendrá su confirmación o negación cuando a) sea probado o en el caso de b) el día 3 de Febrero de 1982 y por tanto carecería de sentido pretender estimar su probabilidad matemática ahora porque, insistimos, se trata de hechos aislados (afirmación b)) o juicios específicos (afirmación a)).

ESPACIO MUESTRAL

Vamos a ocuparnos en primer lugar y de manera exhaustiva de la definición clásica de probabilidad y de su utilidad para resolver determinados tipos de problemas con unas características que en su momento describiremos. Previo a ello hemos de introducir una serie de conceptos preliminares que nos serán útiles posteriormente.

Consideremos un conjunto fijo de condiciones $\mathcal{G}$ y una familia $\mathcal{S}$ de sucesos, A, B, C... que pueden ocurrir o no cuando se realiza $\mathcal{G}$. Existen entre los elementos de $\mathcal{S}$ ciertas relaciones de interés que vamos a poner de manifiesto.

1.- Si para cada realización del conjunto de condiciones $\mathcal{G}$ en la que ocurre el suceso A ocurre también el suceso B, decimos que A <u>implica</u> B y lo denotaremos mediante

$$A \subset B \quad \text{o} \quad B \supset A.$$

2.- El hecho de que A y B se impliquen mutuamente supone que ambos sucesos son <u>equivalentes</u>, lo que denotamos mediante $A = B$. Hemos de señalar que por lo que respecta a la teoría de la probabilidad, los sucesos equivalentes pueden reemplazarse uno por otro y por tanto en adelante los consideraremos simplemente como idénticos.

3.- Un suceso que consiste en la ocurrencia simultánea de otros dos A y B se llamará <u>producto</u> de A y B y lo denotaremos AB (o $A \cap B$).

4.- Un suceso que consiste en la ocurrencia de al menos uno de los sucesos A y B se llama <u>suma</u> de A y B, su representación es $A+B$ (o $A \cup B$).

5.- Un suceso que consiste en el hecho de que ocurra A pero no B se llama <u>diferencia</u>

de A y B y se designa mediante $A-B$.

Ilustremos mediante un ejemplo estos nuevos conceptos. Supongamos que G consiste en el hecho de lanzar un dado una vez. Sea A obtener un seis en el lanzamiento, B obtener un tres, C obtener un número par de puntos y D obtener un múltiplo de 3. Los sucesos A, B, C y D están relacionados de la siguiente manera

$$A \subset C \ , \ A \subset D \ , \ B \subset D$$

$$A + B = D \ , \ CD = A$$

La definición de suma y producto de dos sucesos puede generalizarse a un número finito de sucesos mediante

$$A + B + \dots + N \quad (A \cup B \cup \dots \cup N)$$

que significa un suceso que consiste en la ocurrencia de al menos uno de los sucesos $A, B, \dots, N$. Análogamente

$$AB \dots N \quad (A \cap B \cap \dots \cap N)$$

significa un suceso consistente en la ocurrencia de todos los sucesos $A, B, \dots, N$.

6.- Un suceso se llama <u>cierto</u> si ocurre necesariamente al realizar G. Por ejemplo al lanzar dos dados la suma de sus caras es ciertamente mayor o igual a 2.
Cuando el suceso no ocurre bajo las condiciones utilizadas se le llama suceso <u>imposible</u>. Por ejemplo obtener 13 puntos en el lanzamiento de los dos dados.
Como todos los sucesos ciertos e imposibles son equivalentes entre sí es costumbre designarlos genéricamente mediante las letras Ω y ϕ respectivamente (o cualquier otras, pero siempre la misma para todos ellos).

7.- Dos sucesos A y $\bar{A}$ se llaman <u>contrarios</u> si se verifican simultáneamente las dos relaciones

$$A + \bar{A} = \Omega \quad y \quad A\bar{A} = \phi$$

Por ejemplo, si C representa obtener un número par de puntos al lanzar un dado, entonces

$$\bar{C} = \Omega - C \quad \text{representa obtener una suma impar de puntos.}$$

8.- Dos sucesos A y B se dice que son <u>mutuamente excluyentes</u> o <u>incompatibles</u> si no pueden ocurrir simultáneamente, es decir, si

$$AB = \phi$$

Si

$$A = A_1 + A_2 + \dots + A_n$$

con los A_i incompatibles entre sí, es decir, $A_iA_j = \phi$, $i \neq j$, decimos que A puede descomponerse en sucesos particulares $A_1 \dots A_n$, o simplemente que los $A_1 \dots A_n$ constituyen <u>una partición</u> de A. Por ejemplo, si C es el número par de puntos y C_2, C_4 y C_6 representan obtener 2, 4 o 6 en el lanzamiento respectivamente, entonces $C = C_2 + C_4 + C_6$ y además son incompatibles entre sí.

Los sucesos $B_1, B_2, \dots, B_n$ forman un <u>grupo completo de sucesos</u> si al menos uno de ellos ocurre en la realización de G. Es decir

$$B_1 + B_2 + \dots + B_n = \Omega.$$

Especialmente interesantes son los <u>grupos completos de sucesos incompatibles dos a dos</u> también llamados <u>particiones</u>.

9.- Cualquier problema en teoría de la probabilidad supone la existencia de un conjunto de condiciones G y de una familia de sucesos $\mathcal{S}$ que ocurren o no cuando se realizan las condiciones de G. La experiencia demuestra que es aconsejable hacer las siguientes suposiciones respecto del sistema descrito:

a) Si la familia $\mathcal{S}$ incluye los sucesos A y B, incluye también los sucesos AB, $A+B$, $A-B$.

b) La familia $\mathcal{S}$ contiene a los sucesos cierto e imposible.

Una familia de sucesos que satisface estas condiciones se denomina <u>álgebra de sucesos</u>.

Observamos que en los ejemplos utilizados como ilustración, siempre es posible aislar sucesos que no admiten descomposición en otros más simples: obtener un determinado número de puntos, entre 1 y 6 al lanzar un dado, obtener cara o cruz al lanzar una moneda, etc. A este tipo de sucesos que no admiten descomposición los llamaremos <u>sucesos simples</u> o <u>elementales</u>.

En la construcción de una teoría matemática de las probabilidades parece necesaria una formalización mayor de la llevada a cabo en los párrafos precedentes. En las exposiciones modernas de la teoría de la probabilidad el punto de partida es un conjunto de sucesos elementales al que denominamos espacio de sucesos simples o <u>espacio muestral</u>. La naturaleza de los elementos de dicho espacio no se especifica de antemano de manera que permita la suficiente libertad de elección como para abarcar todas las situaciones posibles. A los subconjuntos del espacio muestral se le conoce como <u>sucesos aleatorios</u>. Un suceso consistente en todos los puntos del espacio muestral es conocido como el <u>suceso cierto</u> o <u>seguro</u>. Todas las relaciones escritas anteriormente entre los sucesos se verifican también cuando llevamos a cabo una construcción formal de la teoría. En su momento volveremos sobre el tema para llevar a cabo de forma exhaustiva dicha construcción. Por ahora continuaremos con la situación descrita en los párrafos anteriores y en particular, para la definición clásica del concepto de probabilidad tendremos en cuenta tan solo <u>los espacios muestrales con un número finito de elementos</u>.

Digamos finalmente que entre los sucesos y las relaciones establecidas se verifican las siguientes leyes, que enunciamos sin demostración:

<u>Ley Conmutativa</u> : $A + B = B + A$, $AB = BA$
<u>Ley asociativa</u> : $A + (B + C) = (A + B) + C$, $A(BC) = (AB)C$
<u>Ley distributiva</u> : $A(B + C) = AB + AC$, $A + (BC) = (A + B)(A + C)$
<u>Ley idempotente</u> : $A + A = A$, $AA = A$

LA DEFINICION CLASICA DE PROBABILIDAD

La definición clásica de probabilidad reduce el concepto de probabilidad a la noción de equiprobabilidad entre los sucesos elementales, noción que se considera básica y no está sujeta a definición formal.

Sea Ω un conjunto de n sucesos elementales todos ellos con igual probabilidad de ocurrir,

$$\Omega = \{E_1, \dots, E_n\}.$$

Designemos por $\mathcal{S}$ la familia formada por Ω, el suceso cierto, ϕ, el suceso imposible, los E_k, $k=1,\dots,n$ y cualquier otro suceso A que admita una descomposición en elementos de Ω. Por ejemplo, si $\Omega = \{E_1, E_2, E_3\}$, entonces $\mathcal{S} = \{\Omega, E_1, E_2, E_3, E_1+E_2, E_1+E_3, E_2+E_3, \phi\}$. Se comprueba fácilmente que la familia $\mathcal{S}$ constituye una álgebra de sucesos. La definición clásica de probabilidad puede formularse como sigue:

Definición: Si A, un suceso de $\mathcal{S}$, está compuesto por m de los sucesos elementales que componen Ω, entonces la probabilidad del suceso A es

$$P(A) = \frac{m}{n}.$$

Esta fórmula es conocida con el nombre de fórmula de Laplace por ser este científico quien la formuló por primera vez en el siglo XVIII. Se enuncia también en ocasiones diciendo que "la probabilidad de un suceso A es el cociente entre el número de resultados del experimento favorables a la realización del mismo y el número de resultados posibles".

Ejemplo.- Supongamos que lanzamos un dado, el espacio muestral estará formado por $\Omega = \{E_1, \dots, E_6\}$ donde E_i representa el hecho de que haya aparecido una cara con i puntos. Definimos

$$G = E_2 + E_4 + E_6$$

que es un suceso que representa el hecho de que haya aparecido un número de puntos par en la cara, de acuerdo con la definición dada

$$P(G) = \frac{3}{6} = \frac{1}{2}.$$

Para cada uno de los sucesos elementales y siempre de acuerdo con la definición $P(E_i) = \frac{1}{6}$, $i=1,\dots,6$

Volvamos nuevamente a la definición de probabilidad que acabamos de dar. Observamos que $P(A)$ puede considerarse como una función definida sobre el álgebra de conjuntos, más exactamente como una función de conjunto definida sobre el álgebra $\mathcal{S}$ de los sucesos. Como tal función y a la vista de su definición posee las siguientes propiedades.

Propiedades de la probabilidad

1) $\forall A \in \mathcal{S} \quad P(A) \geq 0$

2) Para el suceso cierto, Ω, $P(\Omega) = 1$

3) Si A puede descomponerse de la forma $A = B + C$, con $BC = \phi$, entonces

$$P(A) = P(B) + P(C).$$

4) La probabilidad de $\bar{A}$, el suceso contrario de A, viene dada por

$$P(\bar{A}) = 1 - P(A).$$

5) La probabilidad del suceso imposible, ϕ, es cero.

6) Si un suceso A implica un suceso B, entonces

$$P(A) \leq P(B).$$

7) La probabilidad de cualquier suceso está comprendida entre 0 y 1, es decir

$$0 \leq P(A) \leq 1, \quad \forall A \in \mathcal{S}.$$

Las demostraciones de estas propiedades es sencilla y puede hacerse como ejercicio.

EJEMPLOS DE APLICACION DE LA DEFINICION CLASICA DE PROBABILIDAD

Consideraremos a continuación algunos casos en los que el cálculo de las correspondientes probabilidades puede llevarse a cabo mediante la utilización de la definición estudiada, por cuanto las condiciones del problema satisfacen las condiciones exigidas en aquella. Nos ocuparemos también de algunos resultados de combinatoria que son interesantes y útiles para la solución de este tipo de problemas.

EJEMPLO 1.- De una baraja de 36 cartas se extraen 3 al azar. Encontrar la probabilidad de que haya exactamente un as entre ellas.

El espacio muestral, todas las posibles extracciones de 3 cartas, está formado por $\binom{36}{3}$ puntos y además todos ellos son igualmente posibles. El número de estas extracciones que contienen un solo as puede determinarse de la siguiente manera: una de las cartas ha de ser un as y podrá ser uno de los cuatro, podrá ser elegido de $\binom{4}{1}$ maneras distintas. Las otras dos cartas deberán elegirse al azar entre las 32 cartas restantes del mazo y ello podrá hacerse de $\binom{32}{2}$ maneras, en definitiva los casos favorables serán $\binom{4}{1} \cdot \binom{32}{2}$ y la probabilidad del suceso A, hay una carta entre las tres extraídas, será

$$P(A) = \frac{\binom{4}{1}\binom{32}{2}}{\binom{36}{3}} = \frac{496}{1785} \simeq 0.2778.$$

EJEMPLO 2.- De una baraja de 36 cartas se extraen 3 al azar. Probabilidad de que haya al menos un as entre ellas.

Sea A el suceso que nos interesa. Este suceso puede descomponerse en tres 3 incompatibles entre sí, a saber, A_1, hay un solo as en las tres cartas extraídas, A_2, hay 2 ases en las cartas extraídas y A_3, las tres son ases. De acuerdo con una propiedad de la probabilidad tendremos

$$P(A) = P(A_1) + P(A_2) + P(A_3).$$

Razonando como en el ejemplo anterior tendremos

$$P(A_i) = \binom{4}{i}\cdot\binom{32}{3-i}\Big/\binom{36}{3} \qquad , i=1,2,3.$$

y de aquí

$$P(A) = \frac{109}{3.119} \simeq 0.3053.$$

Puede resolverse ste problema de otra forma. En efecto, podemos calcular la probabilidad de $\bar{A}$, suceso que representa el hecho de que no haya ningún as entre las cartas extraídas, los casos favorables se obtienen eliminando los 4 ases del mazo y eligiendo las tres cartas entre las 32 cartas restantes, de aquí

$$P(\bar{A}) = \frac{\binom{32}{3}}{\binom{36}{3}} = \frac{31.8}{3.17.7} \simeq 0.6947$$

y de aquí $P(A) = 1 - P(\bar{A}) \simeq 1 - 0.6947 \simeq 0.3053.$

Estos dos ejemplos ponen de manifiesto la utilidad de la combinatoria en la solución de ste tipo de problemas. Vamos a dar a continuación algunos resultados de combinatoria que serán de gran ayuda.

Algunos resultados de combinatoria.

El primero de stos resultados, que se conoce en ocasiones con el nombre del Principio Fundamental del conteo, s de demostración inmediata y afirma:

Teorema 1.- Sea T un trabajo o tarea que requiere, para llevarse a cabo, de la presencia de ~~un~~ total de k trabajos, T_i, $i=1,\ldots,k$, de manera que cada uno de stos puede realizarse de n_i, $i=1,\ldots,k$ formas distintas. Entonces, el número de formas distintas de ejecutar T viene dado por $\prod_{i=1}^{k} n_i$.

A continuación vamos a considerar problemas de seleccionar bolas de una urna o de colocar bolas en celdas, problemas que sirven como modelos generales para gran número de problemas de la vida real. En todos ellos se aplicará implícita o explícitamente el principio fundamental que acabamos de enunciar.

Consideremos una urna con n bolas numeradas, todas ellas idénticas. Si extraemos k bolas de la urna, diremos que hemos extraído una muestra de tamaño k. La muestra decimos que s ordenada si el orden en que las bolas fueron extraídas s tenido en cuenta, en cualquier otro caso diremos que la muestra s desordenada. Vamos el siguiente resultado.

Teorema 2.- El número de muestras diferentes de tamaño k, ordenadas, es el número de variaciones de n elementos tomados de k en k, $V_{n,k} = n(n-1)\ldots(n-k+1)$, si las extracciones han sido hechas sin reemplazamiento y s igual a n^k si las extracciones se llevaron a cabo con reemplazamiento.

Para muestras desordenadas dichos números son respectivamente

$\binom{n}{k}$ para muestras sin reemplazamiento, y

$\binom{n+k-1}{k}$ para muestras con reemplazamiento.

Demostración.- Para las muestras ordenadas bastará aplicar el principio antes enunciado con $n_j = (n-j+1)$ y $n_j = n$, $j = 1,\ldots,k$, respectivamente.

Para las muestras desordenadas, extraídas sin reemplazamiento, observamos que cada una de ellas da lugar a $k!$ muestras ordenadas, por tanto si su número s x, tendremos

$$V_{n,k} = k!\,x \quad , \text{ de aquí } \quad x = \frac{V_{n,k}}{k!} = \binom{n}{k}$$

para las extracciones con reemplazamiento puede recurrirse al método de inducción para comprobar la fórmula dada, no obstante vamos a presentar un método alternativo más ocurrente. Supongamos que las n bolas son n cartas numeradas del 1 al n y que les añadimos $k-1$ cartas, numeradas del $n+1$ a $n+k-1$ conteniendo cada una una inscripción que dice: "repita la carta menor", "repita la 2ª carta menor", ——, "repita la $k-1$ sima carta menor". Extraer una muestra de tamaño k con reemplazamiento de entre las n cartas iniciales s equivalente a extraer una muestra de tamaño k sin reemplazamiento de entre las $n+k-1$ cartas, así pues su número vendrá dado por $\binom{n+k-1}{k}$.

Una aplicación de stas propiedades podemos verla en el siguiente ejemplo:

Ejemplo.- ¿Cuál s la probabilidad de que una mano de poker contenga exactamente un par?

Una mano de poker es un subconjunto de 5 cartas elegidas entre las 52 del mazo de cartas. El espacio muestral de las posibles manos de poker s finito y contiene exactamente $\binom{52}{5}$ puntos. Por otra parte, una mano de poker con un par contiene dos cartas iguales y las otras tres diferentes entre sí y distintas de las que forman el par. Podemos obtener semejante mano mediante los siguientes pasos:

1) Elegir la carta que formará el par entre las 13 distintas de cada palo. Esto puede hacerse de $\binom{13}{1}$ formas

2) Elegida la carta en 1), elegir las dos cartas de se número entre las cuatro (una de cada palo) que van a constituir la pareja. Lo haremos de $\binom{4}{2}$ formas distintas

3) Elegir ahora las tres cartas restantes, los números correspondientes, entre los 12 números que quedan después de eliminar el correspondiente a la pareja. Hay un total de $\binom{12}{3}$ formas de hacerlo

4) Finalmente determinar para cada uno de los tres números elegidos, cuál s el palo elegido lo podemos hacer de $4.4.4 = 4^3$ formas distintas.

Los casos favorables son pues $\binom{13}{1}\cdot\binom{4}{2}\cdot\binom{12}{3}\cdot 4^3$. En definitiva la probabilidad pedida vendrá dada por

$$\frac{\binom{13}{1}\binom{4}{2}\binom{12}{3}4^3}{\binom{52}{5}} = \frac{1098240}{2598960} \simeq 0.42$$

Otro problema interesante que puede plantearse s el de la ocupación de celdas mediante bolas. Genéricamente ste tipo de problemas son conocidos con el nombre de problemas de ocupación. El teorema siguiente s útil en la resolución de los mismos.

TEOREMA 2.- i) El número de maneras distintas en que n bolas distintas pueden distribuirse en k celdas distintas es k^n.

ii) Si deseamos que cada celda contenga un determinado número de bolas, es decir, la j-ésima celda contenga n_j bolas, $j=1,\ldots,k$ ($n_j\geq 0$, $j=1,\ldots,k$. $\sum n_j = n$), entonces las maneras de llevar a cabo su distribución son

$$\frac{n!}{n_1!\,n_2!\cdots n_k!} = \binom{n}{n_1,n_2,\ldots,n_k}.$$

iii) Si las n bolas son indistinguibles pero las k celdas son distintas, entonces el número de distribuciones es

$$\binom{k+n-1}{n}.$$

Además, si $n\geq k$ y ninguna celda debe de estar vacía, entonces esta cantidad se convierte en

$$\binom{n-1}{k-1}.$$

Demostración.- i) Cada una de las n bolas puede ocupar k posiciones distintas, por tanto k^n es el número de posibles distribuciones.

ii) El problema puede asimilarse al de establecer k grupos con las n bolas de manera que cada uno de ellos contenga un número determinado de bolas, exactamente n_j el j-ésimo grupo. Esto podemos hacerlo como sigue

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\cdots-n_{k-1}}{n_k} = \frac{n!}{n_1!\,n_2!\cdots n_k!} = \binom{n}{n_1,n_2,\ldots,n_k}.$$

iii) El problema puede reducirse a una de las situaciones descritas y resueltas en el teorema anterior. Supongamos una urna con k bolas numeradas de 1 a k, llevamos a cabo n extracciones con reemplazamiento y colocamos una bola en la celda que indica el número de la bola extraída, obviamente el orden de aparición es irrelevante. Bastará pues con determinar el número de muestras desordenadas de tamaño n, extraídas con reemplazamiento, que pueden obtenerse de una urna con k bolas numeradas de 1 a k. De acuerdo con el anterior teorema su número es

$$\binom{k+n-1}{n}$$

Si $n\geq k$ y queremos que todas las celdas contengan una bola al menos, depositamos la bola correspondiente en cada una y distribuimos como antes las $n-k$ restantes, tendremos

$$\binom{k+(n-k)-1}{n-k} = \binom{n-1}{n-k} = \binom{n-1}{k-1}.$$

EJEMPLO 3.- Tenemos n partículas, cada una de las cuales puede ocupar cada una de N ($N>n$) celdas con la misma probabilidad, $1/N$. Encontrar la probabilidad de que:

a) Haya una partícula en cada una de n celdas fijadas de antemano,

b) Haya una partícula en cada una de n celdas elegidas al azar.

Este problema es un problema clásico de ocupación que juega un papel importante en Mecánica Estadística. Como las condiciones del problema no aparecen determinadas de manera única existen varias respuestas al mismo, tantas como distintas maneras de definir el espacio muestral (conjunto de los posibles resultados), que dan lugar a las distintas estadísticas, conocidas en Mecánica por el nombre de los autores que las desarrollaron. Exactamente tres son las respuestas que podemos dar a esta cuestión, son

1) Estadística de Boltzmann. Supongamos que las bolas y las celdas son distinguibles, y haciendo uso del teorema 2, apartado i) los casos posibles vendrán dados por N^n. Las respuestas a las dos probabilidades buscadas son

1.a) $p_1 = \dfrac{n!}{N^n}$, hay $n!$ formas de disponer las partículas distinguibles en las n celdas distintas.

1.b) $p_2 = \dfrac{\binom{N}{n}\,n!}{N^n} = \dfrac{N!}{N^n(N-n)!}$, $\binom{N}{n}$ son las formas de elegir las n celdas entre las N disponibles.

2) Estadística de Bose-Einstein. Si bien las celdas continúan siendo distintas, las n partículas son ahora indistinguibles, por tanto, y recurriendo al teorema 2 el número de casos posibles es ahora $\binom{n+N-1}{n}$, y las probabilidades buscadas

2.a) $p_1 = \dfrac{1}{\binom{n+N-1}{n}}$, los casos favorables son 1 al no distinguirse las partículas

2.b) $p_2 = \dfrac{\binom{N}{n}}{\binom{n+N-1}{n}} = \dfrac{N!\,(N-1)!}{(N-n)!\,(N+n-1)!}$

3) Estadística de Fermi-Dirac. En esta estadística se supone que las celdas pueden contener a lo sumo 1 partícula, partícula que además no se distingue de las otras. Los casos posibles son pues $\binom{N}{n}$ y las probabilidades

3.a) $p_1 = \dfrac{1}{\binom{N}{n}}$

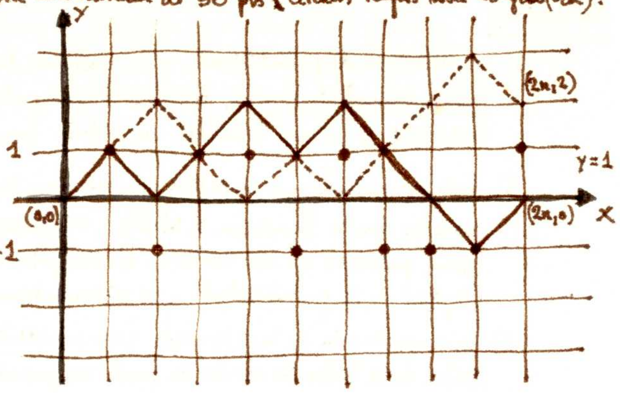
3.b) $p_2 = \dfrac{\binom{N}{n}}{\binom{N}{n}} = 1.$

EJEMPLO 4.- En la taquilla de un matinal de cine hay una cola de $2n$ niños. El precio de la entrada son 50 pesetas y de los $2n$ niños, n llevan monedas de 50 pts y los n restantes llevan billetes de 100 pesetas. Si el cajero no tiene cambio cuando comienza el despacho de entradas, ¿cual es la probabilidad de que ningún niño deba esperar para el cambio?

Resulta dificultoso en este caso la determinación del número de casos favorables por métodos directos, para averiguar tales casos procedemos de la siguiente forma:

Representaremos el orden de llegada a la taquilla en el eje de las x's. Sobre cada uno de los puntos levantaremos una ordenada que valdrá 1 si el niño va a pagar con un billete de 100 pts y -1 si lleva una moneda de 50 pts (Circulos negros sobre la gráfica).

A continuación trazamos una poligonal, partiendo de $(0,0)$, cuyos vértices son aquellos puntos que tienen por abcisa el orden de llegada y por ordenada la acumulación de todas las ordenadas anteriores más la correspondiente al punto en cuestión (vease la figura adjunta). A esta poligonal le llamaremos trayectoria. Dos propiedades interesantes tiene la trayectoria así definida:



a) Hay tantas trayectorias como ordenes distintos de llegada, a saber $\binom{2n}{n}$.

b) Todas las trayectorias comienzan en $(0,0)$ y acaban en $(2n,0)$.

De todas las trayectorias, las que se corresponden con ordenes de llegadas favorables a las condiciones requeridas en el enunciado son aquellas que no cruzan el eje de las x's (hacia las y's positivas). Ello supone que en cada abscisa, es decir a la llegada de cada niño, los que llevaban monedas de 50 pts han sido al menos 1 más que los que llevaban 100 pts para pagar su entrada, es decir, tenemos así garantizada la existencia de cambio para el mero comprador. Vamos a determinar el número de estas trayectorias.

Mediante la recta $y=1$, utilizada como eje de simetría, construiremos nuevas trayectorias a partir de las iniciales. Estas nuevas trayectorias comienzan también en $(0,0)$, coinciden con las antiguas mientras estas no alcanzan la recta $y=1$, cuando esto ocurre la recta hace de eje de simetría de la trayectoria. Esto supone:

a) Las trayectorias que no alcanzan $y=1$ conservan su forma inicial y no dan lugar a nuevas trayectorias ficticias. Observese que estas trayectorias son precisamente aquellas que no cruzan el eje de las x's, es decir, las favorables a la situación pedida.

b) Las demas trayectorias, las desfavorables, dan lugar a nuevas trayectorias (linea discontinua en la figura) que todas finalizan en el punto $(2n,2)$. Ello quiere decir que tienen 2 subidas más que bajadas, es decir, 2 individuos más con billete de 100 pts que los que pagan con monedas de 50 pts. Su número será pues $\binom{2n}{n+1}$.

En definitiva el número de casos favorables vendrá dado por

$$\binom{2n}{n}-\binom{2n}{n+1}$$

y la probabilidad pedida será

$$p=\frac{\binom{2n}{n}-\binom{2n}{n+1}}{\binom{2n}{n}}=\frac{\frac{2n!}{n!\,n!}-\frac{2n!}{(n+1)!\,(n-1)!}}{\frac{2n!}{n!\,n!}}=1-\frac{n!\,n!}{(n+1)!\,(n-1)!}=$$

$$=1-\frac{n}{n+1}=\frac{n+1-n}{n+1}=\frac{1}{n+1}.$$

FRECUENCIA Y PROBABILIDAD

La definición clásica de probabilidad encuentra en algunas situaciones, aquellos problemas que se derivan de las ciencias experimentales, por ejemplo, dificultades insalvables para su aplicación correctamente. Es muy difícil aislar y determinar con exactitud los casos igualmente probables o bien las condiciones de simetría que la equiprobabilidad comporta no se verifican en las situaciones estudiadas (desintegración de un átomo en un intervalo de tiempo dado, nacimiento de un niño, etc.)

Buscando soluciones a estas situaciones, se comprueba que las observaciones de la ocurrencia o no ocurrencia de un suceso A, para un gran número de experimentos repetidos bajo un mismo conjunto de condiciones G, muestran que gran cantidad de fenómenos son de tal naturaleza que el número de dichas ocurrencias (no ocurrencias) del suceso A obedece una ley estable. Concretando, si m es el número de veces que ocurre A en n experiencias independientes (los resultados de unas y otras no se influyen), se encuentra que la relación $\frac{m}{n}$, para n suficientemente grande, es prácticamente constante, siendo sus desviaciones cada vez menos frecuentes a medida que aumenta n.

Esta especie de estabilidad de las frecuencias (es decir, la relación m/n) fue puesta de manifiesto primeramente para fenomenos demográficos y ello desde tiempos muy antiguos. Parece ser que en la China del año 2000 antes de Cristo ya había costumbre de confeccionar censos y a partir de las observaciones que ellos permitían se consideraba que la frecuencia de nacimientos de niños era una cantidad prácticamente constante alrededor de $1/2$. La estabilidad de estos nacimientos es un caso especialmente interesante ~~respecto de esta estabilidad~~, podría decirse que incluso anecdótico. En efecto Laplace en su libro Essai philosophique sur les probabilités relata la siguiente anécdota:

Observaciones llevadas a cabo durante varias decadas en varias ciudades europeas, Londres, San Petersburgo, Berlin y en toda Francia le llevaron a la conclusión que el nacimiento de niños presentaba una frecuencia extraordinariamente coincidente en el tiempo y en el espacio, siendo dicho valor 22/43. Estudió Laplace dichos nacimientos durante un un periodo de 40 años (de 1745 a 1784) para la ciudad de París y encontró que su frecuencia era de 25/49, cantidad ligeramente diferente de la anterior, pero lo suficiente para sorprender a Laplace. Profundizando más en su estudio descubrió que en los censos utilizados figuraban también los niños expósitos y que además las poblaciones de los suburbios de París tenían preferencia por abandonar niños de un solo sexo. En aquella época este fenómeno era bastante corriente y lo suficientemente importante como para alterar los resultados de la encuesta que se estaba llevando a cabo. Corregido el censo de estos niños se encontró para París una frecuencia de nacimientos de niños de 22/43.

Desde la época de Laplace son muchos los experimentos llevados a cabo en este sentido. En la tabla se recogen los resultados de distintas experiencias llevadas a cabo en la determinación de la frecuencia con que aparecía cara en una serie de lanzamientos

Realizó la experiencia	Nº de lanzamientos	Nº de caras	Frecuencia
BUFFON	4040	2028	0.5080
KARL PEARSON	12000	6019	0.5016
KARL PEARSON	24000	12012	0.5005

La tecnología actual permite llevar a cabo de manera sencilla y rápida otras comprobaciones de este hecho experimental (por ejemplo mediante la generación de números aleatorios con un ordenador).

La verificación de este hecho induce a pensar en la existencia de ciertas regularidades en determinados fenómenos, inherentes a los mismos y por tanto independientes del experimentador. El hecho de que para los sucesos a los que es aplicable la definición clásica de probabilidad, la estabilidad de sus frecuencias se sitúe alrededor y muy cerca de dicha probabilidad, nos conduce, de manera lógica, a pensar que algo similar ocurre para sucesos de carácter general. No es pues extraño que decidamos denominar probabilidad del suceso A a la constante numérica alrededor de la cual oscilan las distintas frecuencias observadas para la ocurrencia de A.

Definición: Diremos que el suceso A tiene una probabilidad si posee las siguientes peculiaridades:

a) Es posible, al menos teóricamente, llevar a cabo bajo las mismas condiciones, un número ilimitado de experimentos independientes, en cada uno de los cuales A puede o no ocurrir.

b) Cuando el número de experimentos es suficientemente grande se observa que la frecuencia de aparición de A se desvía muy poco de una constante (generalmente desconocida).

Esta constante desconocida llamada probabilidad estadística de A, para un gran número de experiencias puede aproximarse mediante la frecuencia observada.

De esta definición y teniendo en cuenta las propiedades de las frecuencias se derivan las siguientes propiedades para la probabilidad.

Propiedades de la probabilidad

1) La probabilidad de un suceso que es cierto, es la unidad.

2) La probabilidad del suceso imposible es cero.

3) Si un suceso aleatorio C es la suma de un número finito de sucesos incompatibles dos a dos, $A_1, A_2, \ldots, A_n$ que tienen probabilidad, entonces su probabilidad existe y es igual a la suma de las probabilidades de las componentes, es decir

$$P(C) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Es interesante hacer algunas observaciones referentes a esta definición de probabilidad.

Observaciones

1) La definición de probabilidad que acabamos de dar es descriptiva, no se trata de una definición matemáticamente formal tan solo hemos postulado la existencia de la probabilidad bajo ciertas condiciones y hemos indicado un método para aproximar su valor.

2) Señalar además que la definición no pone de manifiesto las peculiaridades de aquellos fenómenos para los que la frecuencia es estable, si quisiéramos averiguarlas deberíamos llevar a cabo investigaciones posteriores en este sentido. En cualquier caso la definición sí pone de manifiesto el carácter objetivo de la probabilidad y su independencia del investigador que lleva a cabo la experiencia.

3) La importancia de la definición frecuencialista de la probabilidad (también se la conoce con este nombre) estriba en el hecho de servir de punto de partida para la construcción de una teoría formal de la probabilidad que considera a esta como una parte de la Matemáticas. Tanto es así, que en ocasiones la Teoría de la Probabilidad se define "como la parte de las Matemáticas que trata de los modelos matemáticos de los fenómenos aleatorios que poseen la propiedad de la estabilidad de las frecuencias"

CONSTRUCCION AXIOMATICA DE LA TEORIA DE LAS PROBABILIDADES

En las dos definiciones estudiadas hasta ahora ha quedado puesto de manifiesto la limitación de las mismas. Esta limitación es doble, por un lado ambas adolecen de una falta de rigor matemático o de formalización matemática si se prefiere, por otra parte ninguna de las dos es exhaustiva por cuanto dejan fuera de su campo de aplicación muchos fenómenos. Esto, unido a las exigencias cada vez mayores de las ciencias experimentales, como primeras usuarias del concepto, planteó, a principios de siglo la necesidad de dar una solución general al problema. La solución pasaba, desde el punto de vista matemático, por la elaboración de una teoría formal mediante una axiomática adecuada, tal como se hace en otras partes de la Matemática. Los axiomas iniciales no suponían, o debían suponer, un inicio en el desarrollo de la teoría, es decir, un estado inicial sin información previa, sino que por el contrario debían derivarse de los conocimientos que los resultados previos habían proporcionado (véase definiciones clásica y frecuencialista).

Así las cosas S.N. Bernstein llevó a cabo un primer intento, en este sentido, hacia 1917. Construyó una axiomática sobre la base de comparar sucesos aleatorios en función de su mayor o menor probabilidad.

La solución definitiva al problema fue propuesta por A.N. Kolmogorov que definió una axiomática, hacia 1930, que relacionaba la teoría de la probabilidad con la teoría de funciones y la teoría de conjuntos. Partía Kolmogorov de una situación abstracta, para darle la mayor generalidad posible, considerando inicialmente un espacio Ω de <u>sucesos elementales</u> y una familia $\mathcal{S}$ de partes de Ω a cuyos elementos denominaba <u>sucesos aleatorios</u>. La familia de sucesos $\mathcal{S}$ verifica las siguientes condiciones

a) $\Omega \in \mathcal{S}$

b) Si $A \in \mathcal{S}$ entonces $\bar{A} \in \mathcal{S}$

c) Si $\{A_n\}_{n \geq 1} \subset \mathcal{S}$, entonces $\bigcup_{n \geq 1} A_n \in \mathcal{S}$.

Una familia con estas características es conocida con el nombre de <u>σ-álgebra de sucesos</u> o <u>tribu de sucesos</u>.

Los axiomas que introdujo Kolmogorov sobre Ω y los elementos de $\mathcal{S}$ fueron los siguientes:

A.I) A cada elemento de la σ-álgebra $\mathcal{S}$ le asociamos una cantidad no negativa a la que llamaremos su probabilidad. $P(A) \geq 0$, $\forall A \in \mathcal{S}$.

A.II) $P(\Omega) = 1$

A.III) Si $\{A_n\}_{n \geq 1} \subset \mathcal{S}$ y $A_i \cap A_j = \phi$, $i \neq j$, entonces

$$P\Big(\bigcup_{n \geq 1} A_n\Big) = \sum_{n \geq 1} P(A_n).$$

A partir de esta axiomática se desarrolla toda la Teoría Moderna de la Probabilidad, pero no lo haremos así nosotros porque la probabilidad, definida mediante estos axiomas, se inscribe en el contexto más general de la Teoría de la Medida, y así como vamos a actuar en el próximo capítulo.

Antes de terminar con este tema señalar dos propiedades que el sistema de axiomas de Kolmogorov posee:

1) El sistema de axiomas es <u>consistente</u>, puesto que existen objetos reales que satisfacen todos los axiomas. Si tomamos $\Omega = \{a_1, a_2, \ldots, a_n\}$, los a_i cualesquiera y si $\mathcal{S} = \mathcal{P}(\Omega)$, haciendo $P(a_i) = p_i$, con $p_i \geq 0$ y tales que $p_1 + p_2 + \ldots + p_n = 1$ y si para $A \in \mathcal{S}$, tal que $A = \{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\}$, definimos $P(A) = p_{i_1} + \ldots + p_{i_k}$, entonces los axiomas de Kolmogorov se satisfacen.

2) El sistema es <u>incompleto</u> puesto que para un mismo Ω podemos elegir las probabilidades en $\mathcal{S}$ de formas diferentes. Por ejemplo para el caso del lanzamiento de un dado, $\Omega = \{E_1, \ldots, E_6\}$ podemos ~~tomar~~ asignar $P(E_i) = 1/6$, $i = 1, \ldots, 6$, o bien hacerlo de la forma

$$P(E_1) = P(E_2) = P(E_3) = 1/4 \quad y \quad P(E_4) = P(E_5) = P(E_6) = 1/12,$$

por cuanto de ambas formas el sistema de axiomas es satisfecho.

La <u>incompletitud</u> del sistema de axiomas no significa en absoluto su inadecuación para explicar o desarrollar la teoría en estudio, en este caso la teoría de la Probabilidad, al contrario es debido a la esencia de la materia en estudio. En efecto, en el caso del dado, si este es un dado perfecto el modelo probabilístico adecuado para describirlo será el primero, aquel en el que todas las caras tienen probabilidad $1/6$, en caso de tratarse de un dado irregular quizás el segundo modelo se ajuste más a la realidad.