1.2 CONCEPCIONES DE PROBABILIDAD

1.2.1 NATURALEZA DE LA PROBABILIDAD

El vocablo azar proviene del árabe “az-zahr”, que primero significó flor y luego suerte; por asociación, el dado fue denominado azar y puesto que confrontaba a los hombres con la incertidumbre, el término fue usado también para referirse a cualquier hecho fortuito y desfavorable; luego se generalizó para denotar cualquier suceso impredecible. Esta noción evoca una supuesta causa mágica de los sucesos, quizá debidos a la naturaleza o provocados de manera intencional por los dioses.

El azar hizo presencia en el Cosmos con el Cosmos y como parte de él. Luego, por obra del azar, el hombre apareció en el Cosmos y lo ha acompañado durante toda su existencia, para bien o para mal.

“*El hombre tiene mil planes para sí mismo. El azar, sólo uno para cada uno*”
Mencio

Han salido sobrando las explicaciones filosóficas y teológicas que se han dado en torno al azar, porque el azar no tiene que ver nada con la filosofía ni con la teología; el azar se manifiesta al hombre de varias maneras sorpresivas e inesperadas, con hechos gratos o terribles, así como a través del juego.

Como resultado de las condiciones adversas que tuvo que enfrentar en sus orígenes y como un efecto evolutivo de su propia especie, el hombre convirtió lo desconocido en un reto de supervivencia, en una necesidad de jugar, para poder mecerse en el vaivén de la fortuna y la desgracia. Fue la posibilidad de ganar, la que condujo al hombre a inventar juegos de azar, apostando sus bienes contra los bienes de otros hombres, a sabiendas de que si hacía su apuesta contra los designios del azar, perdería.

Se considera que los juegos de azar han estado presentes de una manera u otra y permanentemente, en todas las civilizaciones de la humanidad; su origen es diverso y sería imposible asegurar cuál fue el primero. Los dados han sido, sin duda, uno de los pasatiempos humanos milenarios, siempre presentes a través de la historia, incluido el presente; los ha habido de diversas formas y tamaños, pero los más populares siempre fueron los de seis caras con las esquinas redondeadas, para propiciar que el dado de vueltas y que el resultado sea aún más impredecible; su tecnología es tan antigua o más que la de las escuadras, de modo que la probabilidad es sin duda de mayor edad que la geometría.

La afición al juego ha alcanzado a doctos y a iletrados, a pobres y a ricos, a religiosos y a paganos, sin que en ellos hubiera indicios de reflexión alguna sobre su aritmética. En los dados el resultado está gobernado por la incertidumbre, intérprete de la voluntad divina o del capricho de natura.

Lo que interesa al jugador es cuál será el resultado en el siguiente lanzamiento, y puesto que no es capaz de hacer una predicción certera, solo le queda especular con base en resultados pasados o esperar a que la suerte lo favorezca.

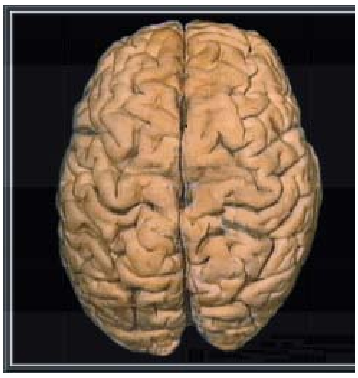
Los jugadores siempre han soñado con poder superar lo fortuito del juego, pues aquel que lograra encontrar un artificio secreto que le permitiera predecir con total exactitud cada siguiente resultado, habría vencido el riesgo y tendría una gran ventaja sobre todos los demás hombres, al apostar sobre seguro a los eventos que sí ocurrirían, obteniendo ganancias ilimitadas sin esfuerzo.

Con esa ilusión fue que el jugador recurrió al matemático, creyendo que podría obtener de él los secretos del juego. Al matemático le fue atractivo el problema, no porque fuera posible dar respuesta al cuestionamiento del jugador, sino por el reto que significaba descubrir los comportamientos del juego; sin embargo, para enfrentar el problema, tuvo que cambiar el enfoque, tomando el siguiente resultado sólo como representativo de resultados hipotéticos futuros y, con ello, pudo formalizar el concepto de probabilidad.

La probabilidad de $1/6$ no le dijo nada al jugador respecto al resultado del siguiente lanzamiento; paradójicamente, al matemático le dijo todo respecto a los resultados de cualquier lanzamiento. Por eso es importante asociar a cada evento un número que mida su posibilidad de ocurrencia, asignando a cada punto muestral del experimento un número llamado medida de probabilidad, o simplemente probabilidad.

Pero la asignación de probabilidades no puede hacerse de manera única, porque los fenómenos estudiados ocurren de distintas maneras y lo que interesa en cada caso puede ser diferente. En ocasiones se puedes suponer que todos los resultados son igualmente posibles, en otras no es válida tal hipótesis; algunos experimentos pueden repetirse indefinidamente en condiciones similares, pero otros son irrepetibles; a veces son relevantes las probabilidades individuales y en otras solo importa el comportamiento probabilístico global. Por ello, la comunidad científica ha ido incorporando las distintas maneras de percibir la probabilidad y las ha aceptado como lícitas, según sea el caso de experimentación de que se trate.

Hasta el Renacimiento, la acepción que se le daba al término probable era algo totalmente ajeno al concepto matemático; probable significaba digno de ser aprobado, haciendo referencia a lo que hacían o pensaban las personas sensatas; de modo que la probabilidad era atributo de opinión, basado en creencias, y tenía como contraparte el conocimiento, que exigía demostración. La dualidad entre conocimiento y creencia se reflejaba como dualidad entre ciencias clásicas y ciencias aplicadas, las primeras generando nuevos conocimientos, y las otras produciendo opiniones basadas en evidencia empírica.



Al surgir la probabilidad como cálculo matemático en el siglo XVII, se realizaron ensayos repetidos de juegos de azar, que mostraron una tendencia a producir frecuencias relativas estables, similares a las probabilidades a priori, por lo que las evidencias experimentales empezaron a ser respetadas, y muy pronto, la probabilidad fue aceptada como disciplina matemática. Desde entonces, las ciencias clásicas siguen deduciendo efectos a partir de causas, mientras que la probabilidad empezó a inferir causas a partir de efectos.

Pascal fue el primero en aplicar un razonamiento probabilístico a un asunto ajeno a los juegos de azar, inventando la teoría de decisiones, al apostar por la existencia de Dios, aun cuando la probabilidad de que exista sea pequeña, porque la ganancia es infinita si en efecto, Él existe. Aunque Pascal haría esta consideración en calidad de misionero, sin proponérselo estableció un camino para incorporar la probabilidad al sentido común, dando la pauta para transferir la estructura de pensamiento de los juegos de azar a una teoría inductiva basada en creencias. La probabilidad había adquirido también una concepción dual: simultáneamente frecuentista y subjetiva.

A raíz del establecimiento de la probabilidad como disciplina matemática empezaron a surgir paradojas que testifican la confrontación entre los enfoques aleatorio e intuitivo. Las diversas interpretaciones de probabilidad entraron en conflicto, porque cada una se quiso formalizar, a costa de invalidar a las otras. La idea de concentrarse en estudiar la estructura de la probabilidad, constituyó un punto de inflexión para su desarrollo y expansión; sin embargo, la teoría axiomática de Kolmogorov, aunque la dotó de una estructura sólida y dio cabida a todas sus interpretaciones, no fue capaz de clarificar su naturaleza.

Las distintas interpretaciones de probabilidad se pueden clasificar en tres grandes categorías: clásica, frecuentista y subjetivista. En el contexto de la ingeniería, cada una es valiosa en sí misma y las tres son aplicables en diferentes situaciones. Para comprender intuitivamente las diferentes maneras de medir las posibilidades de ocurrencia, basta colocarse en cuatro diferentes puntos de vista típicos: jugador, estadístico, experto y neófito.

El jugador sólo requiere conocer bien las condiciones del juego, para poder asignar probabilidades y hacer su apuesta, y no necesita corroborarlas experimentalmente; le interesa que los dados no estén cargados y que cada tirada sea efectivamente al azar, porque así asegura que cada resultado sea igualmente verosímil. Su fórmula mágica es la relación modalidades favorables entre modalidades posibles, y en sentido predictivo, cada probabilidad tiene para él un significado particular, pues él siempre colocará su apuesta por el evento que sabe más probable, así se trate de la primera o de la última tirada.

El estadístico puede asignar probabilidades solo después de haber realizado muchas veces el experimento; se interesa en eventos cuyas causas parecen diversas y variables, comprobando que existe una proporción casi constante entre el número de ocurrencias del evento y el número total de observaciones, porque así asegura una cierta regularidad en los resultados. Su fórmula mágica es la relación número de ocurrencias del evento entre número de observaciones realizadas, y en un sentido predictivo, la probabilidad tiene únicamente un significado global; los datos individuales carecen de relevancia.

El experto, que si bien carece de datos estadísticos, dispone de un acervo de conocimientos que le permiten establecer posibles relaciones lógicas entre causas y efectos. Para asignar probabilidades él requiere conocer a detalle las circunstancias particulares de cada caso, las cuales compara mentalmente con comportamientos pasados similares y pondera las evidencias presentes, con base en su experiencia; cada posible resultado constituye una hipótesis valorada conforme a sus convicciones. Le interesa que no se omita ninguna hipó-

tesis, por ligera que parezca, porque así asegura que todos y cada uno de los posibles resultados puedan ser valorados. Su fórmula mágica es la proporción expresada en términos porcentuales y, en sentido predictivo, por lo general, él podrá estar interesado tanto en el comportamiento probabilístico global como en la probabilidad de un evento específico individual.

El neófito no cuenta ni con información estadística, ni con el conocimiento necesario para establecer relaciones causa-efecto, por lo que su asignación de probabilidades es totalmente subjetiva, respondiendo a su intuición, deseos y temores; por ello, su asignación pueden diferir radicalmente a la dada por otro sujeto. En estos casos, lo normal es que se trate de asuntos intrascendentes; la fórmula mágica también es la proporción porcentual, dada en términos muy burdos y, en cuanto al sentido predictivo él estará interesado en el resultado particular, para efectos de apuesta.

1.2.2 PROBABILIDAD CLÁSICA

Si un evento A contenido en el espacio muestral Ω de un experimento, está formado por $N(A)$ puntos muestrales y el espacio muestral por N puntos muestrales igualmente verosímiles, se dice que la probabilidad de que el evento A ocurra, está dada por la relación:

$$P(A) = \frac{N(A)}{N} = \frac{\text{número de casos favorables}}{\text{número de casos posibles}} \quad \text{(1.13)}$$

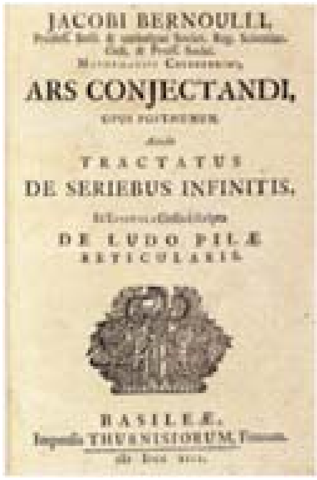
Este es el criterio conocido como interpretación clásica de probabilidad; se dice que es una probabilidad a priori, porque se asigna sin depender de ninguna experiencia, previo a la realización del experimento, o sin que éste se realice; sólo se deben observar dos requisitos: que el número de casos posibles sea finito, y que todos y cada uno de esos casos posibles sean igualmente verosímiles. Es suficiente con que el espacio muestral sea finito para que el primer requisito se cumpla; la satisfacción del segundo es plausible sólo para los juegos de azar, al margen de éstos, es extremadamente difícil que todos los resultados sean igualmente posibles.

La teoría clásica conformó el concepto de probabilidad a partir de un pensamiento determinista. Ante la falta de conocimiento en relación al experimento aleatorio del que se derivan los eventos, supone que existe una simetría recíproca entre los distintos resultados posibles, simetría que permite considerar que todos los resultados posibles pueden ser considerados equivalentes desde el punto de vista de la probabilidad. Por consiguiente, no es cierto que una visión determinista del mundo excluya el pensamiento probabilístico.

Desde la perspectiva matemática, la teoría laplaciana no presenta contradicción alguna; cumple con las exigencias del rigor matemático, a pesar de que no fue construida sobre una base axiomática. Por cuanto a la definición en sí, ésta resulta tautológica, ya que en ella se hace alusión al concepto que se pretende definir; cuando se habla de casos posibles, ello significa igualmente posibles, lo cual equivale a suponer equiprobabilidad.

El criterio clásico fue el primero que utilizaron los estudiosos de los juegos de azar. En 1565 Gerolamo Cardano escribió que la probabilidad consistía en determinar la cantidad de resultados que puede arrojar un juego de azar y determinar también el resultado que favorece al jugador, por lo que la probabilidad es una relación proporcional que se establece entre el resultado favorable al jugador sobre la cantidad de resultados posibles que arroja el juego.

El primero en dar la definición clásica de probabilidad fue Jacob Bernoulli, en su obra "*Ars Conjectandi*", publicada póstumamente en 1713; atraído por los trabajos estadísticos de Graunt y Petty, discurrió que el modelo de probabilidad ideal propuesto por Fermat, Pascal y Huygens, relativos al comportamiento de los juegos de azar, era frecuentista, pues al realizar un determinado número de jugadas, cada posible resultado del juego, al ser equiprobable, debía observar una frecuencia de aparición, acorde a su respectiva probabilidad de ocurrencia, y se podía calcular por anticipado la cantidad esperada de ocurrencias de cada evento.


JACOBI BERNOULLI,
ARS CONJECTANDI,
TRACTATUS
DE SERIEBUS INFINITIS,
DE LUDO PILÆ
RETICULARIS.
BASILEÆ,
Impensis THURNISIORUM, Fratrum.

La interpretación clásica también se conoce como probabilidad laplaciana, porque fue Pierre Simon Laplace quien la estableció formalmente, en su gran obra "*Théorie analityque* des *Probabilités*", publicada en 1812. Equivocadamente, hizo extensiva la aplicación de la definición clásica de probabilidad a toda situación, considerando que todos los fenómenos observables son comparables a un juego de azar, en cuanto a la simetría mutua de sus eventos elementales, haciéndolos igualmente verosímiles, y sin considerar la posibilidad de que el total de posibles resultados fuese infinito. Caracterizaba los eventos como equiprobables invocando el principio de la razón insuficiente, la cual establece que, si no es posible asignar probabilidades a los estados, entonces éstos pueden ser considerados como igualmente probables, y con fundamento en la filosofía de que la naturaleza se supone indiferente, no hay razón establecida para que un estado de la naturaleza sea más probable que cualquiera otro. En aparente contrasentido, Laplace también proyectó, de manera muy influyente, la imagen de un mundo completamente determinista, basado en el principio de causalidad, que postula que todo efecto debe tener siempre una causa; era el más firme creyente del determinismo causal, afirmando que si hubiese un intelecto que conociera el estado actual del mundo con total precisión, podría predecir cualquier evento futuro.

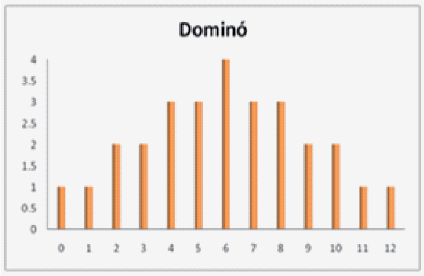
Ejemplo 1.15. DADOS. Consideremos el experimento consistente en lanzar un dado y observar la cara que queda hacia arriba; el espacio muestral asociado es: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Suponemos que el dado no está cargado y que, por lo tanto, no hay ninguna razón para sospechar que alguna cara del dado se puede presentar con mayor facilidad que otra. El criterio de Laplace es el apropiado para asignar probabilidad a cada uno de los seis puntos muestrales o posibles resultados del experimento, porque es un juego de azar con un número finito de posibles resultados, considerados equiprobables.

La probabilidad de que la cara que queda hacia arriba sea el 2 es: $P(2) = \frac{N(2)}{N} = \frac{1}{6}$ Y de manera similar: $P(1) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$

La probabilidad de que al lanzar un dado la cara que quede hacia arriba sea un número par es: $P(par) = \frac{N(par)}{N} = \frac{3}{6} = \frac{1}{2}$

Ejemplo 1.16. DOMINO Y DADOS. Al descubrir al azar una pieza de dominó, la probabilidad de cualquiera de ellas es *1/28*, porque hay *28* fichas, todas ellas distintas; pero si el experimento consiste en observar la suma de puntos, entonces ya no es igual de fácil, por ejemplo, sacar un *6* que un *3*; de hecho la probabilidad de un *6* es el doble de la probabilidad de un *3*, no porque el *6* sea el doble que el *3*, sino porque hay cuatro fichas que suman *6* y solo dos que suman *3*. Si quisiéramos mantener el mismo espacio muestral con *28* posibles resultados, las respectivas probabilidades serían *4/28* y *2/28*; pero si cambiamos el espacio muestral conforme al experimento redefinido, entonces, las probabilidades de *4/28* y *2/28*, corresponden a eventos que ya no son igualmente posibles, porque ya no son eventos elementales.

Al lanzar un dado, la probabilidad de que caiga un *3* o un *6* es la misma: *1/6*, un caso favorable entre seis totales. Y si se lanzan dos dados y se observa la suma de las caras que quedan hacia arriba, la probabilidad de un *3* es de *1/18*, porque hay dos casos favorables: (*1, 2*) y (*2, 1*) de entre *36* totales, y la probabilidad de un *6* es *5/36*, porque hay cinco casos favorables: (*1, 5*), (*2, 4*), (*3, 3*), (*4, 2*) y (*5, 1*) de entre *36* posibles.



En ningún caso puede darse por buena la hipótesis de que todos los resultados del experimento son igualmente posibles. Para asumir equiprobabilidad se requiere tener certeza, a partir de razonamientos lógicos sólidos o evidencias empíricas. Si bien es cierto que en los juegos de azar, los eventos elementales son equiprobables, no son a éstos a los que hay que calcularles su probabilidad de ocurrencia, sino a eventos compuestos, asociados a experimentos redefinidos, que generan espacios muestrales totalmente distintos del original.

Al intentar aplicar la definición clásica fuera del contexto de los juegos de azar, surgen dificultades por cuanto a la determinación de los datos numéricos, la interpretación empírica y la verificación de los resultados. La definición clásica resulta tan endeble que basta con que el resultado se sesgue ligeramente a favor de un resultado, por ejemplo que en un nacimiento sea más factible un niño que una niña, para que la definición ya no sea aplicable, porque deja de haber simetría.

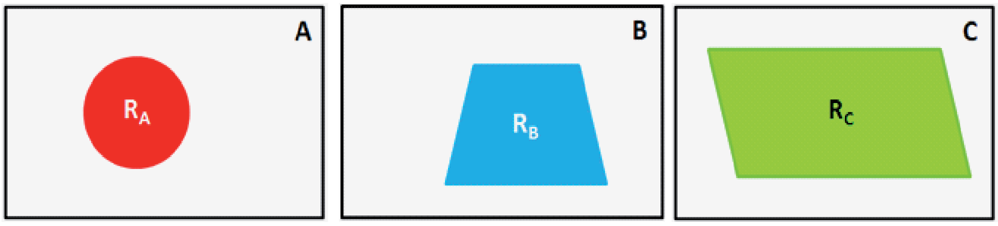
En los juegos de azar la interpretación laplaciana es muy clara, porque los espacios muestrales siempre deben ser finitos; cada resultado elemental es igualmente posible, ya sea que se trate de elegir un objeto al azar, extraer aleatoriamente una bola, o mirar el resultado del lanzamiento de un dado o una moneda.

La dificultad en el cálculo de probabilidades no está en obtener el cociente, sino en contar los casos favorables y los casos totales. La importancia de aprender a contar de manera sistemática, obliga una digresión sobre análisis combinatorio, al cual está dedicado todo el capítulo 1.2.

Otro inconveniente de la definición clásica es que no considera la posibilidad de que el número total de posibles resultados sea infinito. Paradójicamente existen muchos casos de aplicación, en los que es evidente la equiprobabilidad en espacios muestrales continuos.

Probabilidad geométrica

Consideramos tres rectángulos A, B y C, con idénticas dimensiones; en su interior hay un círculo, un trapecio y un paralelogramo, respectivamente, cuyas áreas se aprecian visualmente y son fáciles de calcular.



El experimento aleatorio consistente en elegir un punto al azar en el interior de cada uno de los rectángulos A, B y C, y se desea calcular las probabilidades de que los puntos elegidos pertenezcan a las regiones R_A, R_B y R_C contenidas en ellos.

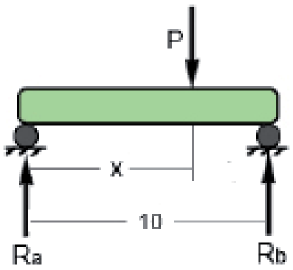
El problema planteado tiene sentido porque las áreas de las regiones coloreadas son comparables, siendo obviamente más fácil acertar a la región R_C que a la región R_B y más fácil acertar a la región R_B que a la región R_A.

Por supuesto que la definición clásica no parece tener lugar en este tipo específico de problemas, pues el cociente $N(R_A)/N$ resulta en una indeterminación matemática, al haber un número infinito de puntos muestrales favorables y un número infinito de puntos muestrales posibles. Sin embargo, aunque la manera de hacer los cálculos no parece tan obvia, el criterio de Laplace se hace aplicable, con una simple modificación:

$$P(R_i) = \frac{\text{área favorable a } R_i}{\text{área total posible}}, \quad i = A,B,C$$

Así, cuando el espacio muestral corresponda a una magnitud continua de longitud, ángulo, área o volumen, el numerador y el denominador del cociente $N(A)/N$ se tienen que traducir a longitudes, ángulos, áreas o volúmenes, respectivamente.

El cálculo de las probabilidades de esas regiones, realizado de esta manera, implica considerar tales regiones como eventos y si éstos se consideran contenidos en un mismo espacio muestral Ω, sobreponiendo los tres rectángulos, entonces con esos eventos se pueden realizar todo tipo de operaciones de unión, intersección, diferencia y complemento. De igual manera, las probabilidades de esos eventos se pueden manipular, conforme a las leyes y reglas de la teoría de la probabilidad, en forma estándar.



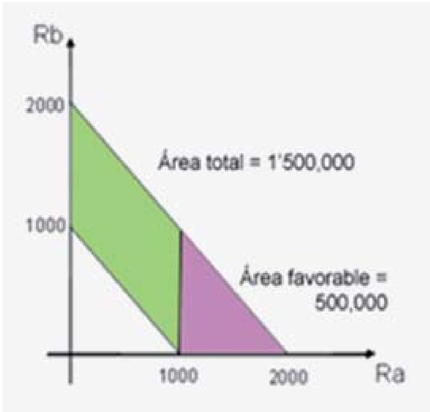
Ejemplo 1.17. VIGA. Considere la viga libremente apoyada y suponga que se desea calcular la probabilidad de que la reacción en el apoyo a sea mayor de *1000* kg, es decir, la probabilidad del evento $A=\{(R_a,R_b)\,|\,R_a\geq 1000\}$

Se supone que los valores de la carga P son equiprobables en el intervalo [*1000, 2000*] y que todos los valores de la distancia x, de la carga P al apoyo a también son equiprobables, en el intervalo [*0, 10*].

La primera dificultad proviene de que el espacio muestral es continuo y por lo tanto el número de casos posibles N es infinito: $R_a\subset[0,2000]$; una segunda dificultad resulta de que el espacio muestral es bidimensional y también se debe tomar en cuenta la reacción R_b:

$$A=\{(R_a,R_b)\,|\,1000\leq R_a\leq 2000, 0\leq R_b\leq 1000, 1000\leq R_a+R_b\leq 2000\}.$$

$$P(A)=\frac{\text{área favorable}}{\text{área total posible}}=\frac{500{,}000}{1'500{,}000}=\frac{1}{3}$$



En 1692, John Arbuthnot publicó la traducción al inglés del "*De ratiociniis en ludo aleae*" de Huygens, al que añadió, entre otras cosas, el siguiente problema: "se lanza un cubo, con aristas a, b, c que se cortan en ángulo recto; ¿cuál es la probabilidad de que el cubo caiga sobre el plano que contiene a dos de esas aristas y es perpendicular a la otra?". Era el primer problema de probabilidad geométrica, diferente a todos los planteados antes, y para su solución, había que identificar casos favorables y casos posibles, en términos de áreas.

Desde entonces, la probabilidad geométrica empezó a desempeñar un papel relevante para el establecimiento de fundamentos de la teoría de la probabilidad, por el reto que representaba poder resolverlos y la polémica que despertaba la aparición de nuevas paradojas.

Durante el siglo XIX hubo varios estudiosos atraídos por la probabilidad geométrica, destacando Joseph Bertrand, quien en 1888, en su obra "*Calcul des probabilités*" introdujo un ejemplo que demostraba que las probabilidades pueden no estar bien definidas; ejemplo que pasó a la historia, conocido como la paradoja de Bertrand.

Fue ya en el siglo XX cuando, motivados por su gran aplicabilidad a la física y a la ingeniería, los seguidores de la probabilidad geométrica crecieron en todo el mundo.

1.2.3 PROBABILIDAD FRECUENTISTA

Si un evento A asociado a un experimento que se ha repetido n veces, se ha observado ese evento en n_A ocasiones, el cociente $f_A = n_A /n$ se llama frecuencia relativa del evento A, la cual satisface las siguientes propiedades:

- La probabilidad es una proporción que se expresa mediante una fracción: n_A /n, en la que los resultados favorables son menores o cuando más iguales a la cantidad de resultados posibles que arroja el juego: $n_A \leq n$
- El valor de esa fracción siempre estará ubicado entre 0 y 1: $0 \leq f_A \leq 1$
- Los valores más cercanos a 1 indican una alta probabilidad de ocurrencia de un resultado. Si $f_A = 1 \Leftrightarrow A$ ocurrió en las n repeticiones.
- Los valores más cercanos a 0 indican una baja o nula probabilidad de ocurrencia. Si $f_A = 0 \Leftrightarrow A$ no ocurrió en ninguna de las n repeticiones.
- Si dos eventos son mutuamente exclusivos, la frecuencia relativa de al menos uno de ellos es la suma de sus frecuencias relativas:

$$A \cap B = \phi \iff f_{A \cup B} = f_A + f_B$$

Como resultado de la observación del comportamiento de numerosas repeticiones de experimentos reales, los estudiosos de la probabilidad se dieron cuenta de que cuando existen razones lógicas para asumir equiprobabilidad, tras una larga serie de realizaciones del experimento, la frecuencia relativa observada de un evento se aproxima a la probabilidad del evento; es decir, la probabilidad a priori, puede ser corroborada a posteriori, empíricamente. Es evidente que, cuando los resultados del experimento no son igualmente posibles, si resulta factible realizar muchas repeticiones del experimento, la frecuencia relativa correspondiente a un evento, puede ser considerada como una buena aproximación de la probabilidad de ese evento.

Si la probabilidad de un evento simple A de un espacio muestral asociado a un experimento tiene un valor $P(A)$ conocido o desconocido, al realizarse repetidamente el experimento, una y otra vez, n veces; si se cuenta el número de veces n_A que el evento A fue observado, la relación n_A/n es tan cercana al valor $P(A)$, que puede considerarse como una buena aproximación del mismo:

$$P(A) \doteq \frac{n(A)}{n} = \frac{\text{número de veces que ocurrió } A}{\text{número de veces que se realizó el experimento}} \qquad (1.14)$$

En muchos problemas, la probabilidad de obtener algún resultado específico se puede interpretar como la frecuencia relativa con la que se obtendría ese resultado, si el experimento se repitiera un número grande de veces, en condiciones similares. No hay claridad de cuál podría ser un número suficientemente grande de veces, ni de qué se debe entender por condiciones similares, pues desde luego no se trata de condiciones idénticas, ya que entonces el experimento ya no sería aleatorio, sino determinista. Tales condiciones son demasiado vagas para servir como base de una definición rigurosa de probabilidad

Por otra parte, si se realizaran otras secuencias de ensayos del mismo experimento, para un evento en particular se obtendrían diferentes valores de su frecuencia relativa, quizá muy similares entre sí, pero finalmente diferentes; no sabríamos cuál de ellos es el más cercano a la verdadera probabilidad.

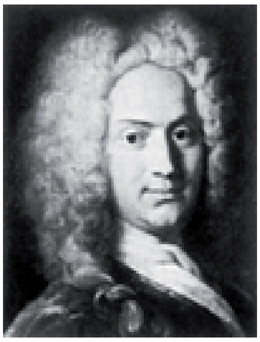
Cuando usamos la frecuencia relativa f_A como aproximación de un valor postulado de la probabilidad $P(A)$, tenemos consciencia de que $f_A \neq P(A)$ y de que estamos cometiendo un error de estimación. Podemos intuir esa propiedad de los fenómenos aleatorios, conocida como regularidad estadística, que consiste en que, a medida que se aumenta el número de repeticiones del experimento, en condiciones similares, varía cada vez menos la frecuencia relativa f_A y tiende a estabilizarse en un valor fijo $P(A)$, o dicho de otra manera, la frecuencia relativa f_A converge a la probabilidad $P(A)$), excepto que esta convergencia en probabilidad debe ser entendida como un hecho empírico y no como un resultado matemático. La estabilidad de las frecuencias en ensayos repetidos es un hecho objetivo de naturaleza independiente del conocimiento de cualquier persona sobre ello.

La interpretación frecuencial de probabilidad consiste en asignar como probabilidad del evento A la relación n_A /n, que es su frecuencia relativa. La probabilidad concebida como proporción hace que la interpretación frecuencial sea fácilmente asimilada. Establecida como criterio formal de asignación, se trata de una interpretación a posteriori, pues la probabilidad solo puede asignarse luego de haber realizado el experimento en forma repetida. Es de esperar que, mientras mayor sea el número de veces que se realice el experimento, esta aproximación será mejor y, en el límite, se obtendrá el valor preciso:

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n} \qquad \text{\_\_\_\_ (1.14')}$$

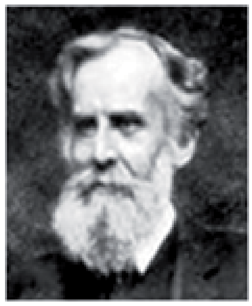
Este resultado es atribuido a Jacob Bernoulli y es conocido como ley de los grandes números, en su forma más básica e intuitiva: la frecuencia relativa tiende a variar cada vez menos, a medida que el número de observaciones va creciendo, estabilizándose alrededor de un valor definido.

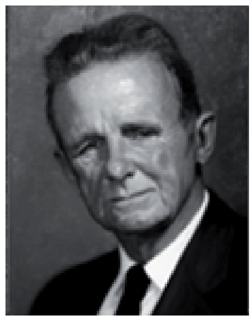
El principio de estabilidad de las frecuencias permite obtener la probabilidad de un evento siempre que se dispone de un número suficiente de observaciones para calcular la frecuencia relativa. Cuando una muestra es tomada correctamente, la frecuencia relativa de un evento supone la probabilidad de dicho evento en el total de la población.



La mayor parte de la gente cree que, en cada nacimiento, es igualmente probable un niño que una niña, y no es así; desde principios del siglo XVIII, Nicolaus Bernoulli aplicó la ley de los grandes números, formulada por su tío Jacob, al registro de *14,000* nacimientos, llegando a la inesperada conclusión de que la frecuencia de nacimientos de niños es mayor que la de niñas, en proporción de *18:17*. Este resultado sigue siendo vigente en el siglo XXI.

La concepción frecuentista de probabilidad consiste en definir la probabilidad como una frecuencia relativa ideal, es decir, como el límite de las frecuencias relativas de un evento, cuando se aumenta indefinidamente el número de realizaciones del experimento. Puesto que es imposible realizar el experimento un número infinito de veces, lo que sí se puede hacer es repetirlo muchas veces, observando que las frecuencias relativas tienden a estabilizarse.





Jacob Bernoulli fue el primero en introducir el criterio de la probabilidad frecuentista o estadística, con objeto de poder calcular la probabilidad de ocurrencia de eventos cuyo número de casos favorables es imposible de contar: “Aquí hay otro camino disponible para alcanzar el resultado deseado; lo que no se puede hallar a priori se puede obtener a posteriori, es decir, mediante la observación múltiple de los resultados de pruebas similares…”

John Venn, el destacado investigador en lógica inductiva, famoso por sus diagramas para representar proposiciones gráficamente, fue el primero en dar un tratamiento sistemático de la concepción frecuentista de probabilidad. En su “The Logic of Chance”, de 1866, estableció el ahora conocido como teorema de Venn: “Si un suceso ocurre en gran número de veces, entonces la probabilidad de ocurrencia del suceso es el límite cuando el número de pruebas tiende a infinito del cociente entre el número de veces que se presenta el suceso y el número total de pruebas” Se trata de un límite empírico que establece la convergencia de la frecuencia relativa a la probabilidad.

En 1919, Richard Von Mises formalizó la definición de la probabilidad frecuencial tal y como se conoce hoy en día; su idea de probabilidad se basó en el valor límite de la frecuencia relativa de un evento en una sucesión indefinidamente prolongada de observaciones, a la que denominó colectivo; cada colectivo conduce a un resultado numérico que satisface dos axiomas: uno de convergencia, referido al límite, y otro de aleatoriedad, que sólo admite colectivos irregulares e impide hacer predicciones, pero asegura que el valor del límite no cambie, cuando se pasa de un colectivo a otro, mediante una elección aleatoria. Las críticas a esta teoría se debían a la combinación de esos dos axiomas, pues no resultaba lógico obtener un límite de una sucesión que no tiene regla alguna.

Para lograr la consistencia de la teoría, en 1935 Hans Reichenbach propuso debilitar la condición de irregularidad al mínimo necesario. Argumentaba que la esencia misma del conocimiento es su incertidumbre, en vista de que las predicciones físicas nunca podrán ser exactas, por la imposibilidad de incorporar todos los factores relevantes en los cálculos, dada la forma como el universo se relaciona con nuestras observaciones, y puesto que los fenómenos empíricos que estudian los científicos son secuencias de eventos que se desenvuelven dentro de ciertos rangos de probabilidad.

La base empírica de los frecuentistas es la regularidad estadística, denominada así por Harald Cramer quien hacia 1950 expresó: “A pesar del comportamiento irregular de los resultados individuales, los resultados promedio, en largas sucesiones de experimentos aleatorios, muestran una sorprendente regularidad”

El criterio frecuentista parece no contraponerse con el criterio clásico, que presupone equiprobabilidad, pero al confrontar los valores de probabilidad de ambos criterios, las diferencias nos pueden llevar a concluir que los eventos elementales no son igualmente posibles, como parecían, o que la frecuencia relativa no aproxima suficientemente bien a la probabilidad postulada.

Ejemplo 1.18. MONEDA. Si se dispone de una moneda ideal, perfectamente equilibrada, al aplicar el criterio de Laplace resulta claro que la probabilidad de obtener un águila cuando la moneda es lanzada es ½; Esta probabilidad es considerada una propiedad de la moneda, como lo es su masa o el material del que está hecha. Con base en esto, se esperaría que la frecuencia relativa de águilas fuera aproximadamente ½, es decir, que la proporción de lanzamientos en los que se obtendría un águila fuese cercana a ½

Las condiciones en las cuales se lanza la moneda no deben ser absolutamente idénticas en cada lanzamiento, porque entonces se obtendrían sólo águilas o sólo soles; las características deseables de aleatoriedad se perderían, al realizar lanzamientos completamente controlados. Investigaciones recientes de la Universidad de Columbia, en Vancouver, señalan que una persona bien entrenada ya es capaz de lanzar una moneda repetidamente de tal manera que puede orientar los resultados en proporciones significativas.

Aunque no se especifica un límite para la variación posible respecto al valor ½, al lanzar la moneda *10* veces, *100* veces, *1000* veces, *10000* veces, *100000* veces o *1000000* de veces, no cabría esperar, en ningún caso, una frecuencia relativa de ½ exactamente, aunque no es imposible que suceda. En el otro extremo, tampoco esperaríamos que tal frecuencia relativa difiriera mucho de ½; si bien puede haber sorpresas para la serie de *10* y de *100* lanzamientos, a medida que es mayor el número de lanzamientos, generalmente ocurre que la frecuencia relativa se va acercando cada vez más a ½.

En una simulación en computadora, el comportamiento de la frecuencia relativa del evento A es muy ilustrativo, pues al compararla con la $P(A) = 0.5$, la inmensa mayoría de las veces será diferente.

n	Sim. 1	Sim. 2
10	0.6	0.4
100	0.48	0.47
1000	0.498	0.5
10000	0.4988	0.4986
100000	0.50046	0.49995
1000000	0.501299	0.498701

La interpretación frecuencial no sirve para asignar probabilidades a eventos aislados que no son susceptibles de experimentación. Los frecuentistas hablan de probabilidades sólo cuando se trata de experimentos aleatorios bien definidos, de modo que este criterio está limitado a los casos en que los resultados presentan regularidad estadística, consecuencia de la observación de un experimento que es considerado aleatorio, porque es repetible bajo las mismas condiciones.

Ejemplo 1.19. PRESA. Considere el espacio muestral del experimento consistente en observar el tirante de agua de una presa cuya cortina tiene una altura de *100* m: $S=[0,100]$ y se desea conocer la probabilidad del evento $B=\{x \mid x \le 30\}$, de que tal nivel sea menor o igual a *30* m. Sin saber cuál es el comportamiento de la presa, no es posible suponer a priori, equiprobabilidad para todos los niveles; la observación de ésta, una y otra vez, por ejemplo durante *1000* días a una misma hora y contar, cuántas de esas veces, el nivel no superó los *30* m; suponiendo que fueron *118* veces. Entonces aplicando el criterio frecuencial de probabilidad, la probabilidad buscada es:

$$P(B)=\frac{n(B)}{n}=\frac{118}{1000}=0.118$$

Ejemplo 1.20. MOTORES. Considere nuevamente el experimento consistente en observar el tablero del avión de cuatro motores. ¿Cuál es la probabilidad del punto muestral (*1, 0, 1, 0*), por ejemplo? Sabemos que este punto muestral indica que fallan dos motores, el primero y el tercero, y que el segundo y el cuarto no fallan. Es necesario primero conocer la probabilidad de falla de cada motor, así como la probabilidad de no falla. Los motores se diseñan para que no fallen y sin embargo fallan; cabe esperar que la probabilidad de falla sea pequeña, pero ¿qué tan pequeña?, ¿cómo asignar esa probabilidad de falla? Suponiendo que los cuatro motores son idénticos, se recurre al experimento consistente en observar el funcionamiento de este tipo de motores. El espacio muestral de este nuevo experimento está constituido por solo dos puntos muestrales: F, falla y F^c, no falla, los cuales obviamente no son equiprobables y para los que, por tanto, el criterio de Laplace no es aplicable.

Supongamos que se observaron *5000* vuelos, bajo condiciones similares, lo que significa la observación de *20,000* motores (*4* por vuelo). Supongamos que se observaron solo dos fallas; entonces, la probabilidad de que un motor falle es: $P(F)=\frac{2}{20000}=0.0001$ y la probabilidad de que no falle es: $P(F^c)=0.9999$

Queda por asignarle probabilidad a los puntos muestrales (*0, 0, 0, 0*), (*1, 0, 0, 0*), etc., pertenecientes al espacio de eventos del experimento original. Pospondremos la solución de este problema para más adelante; por lo pronto conviene distinguir que estos puntos muestrales no son equiprobables, pues si la probabilidad de falla de un motor es muy pequeña, es mucho más fácil que no falle ninguno en un avión de cuatro motores: (*0, 0, 0, 0*), a que falle uno, por ejemplo el segundo: *(0, 1, 0, 0)*

Método de Montecarlo

Es un método estadístico numérico utilizado para aproximar expresiones matemáticas complejas y difíciles de evaluar con exactitud; creado inicialmente para resolver integrales que no tenían solución analítica, luego su uso se extendió a la obtención de soluciones aproximadas de una gran variedad de problemas matemáticos, haciendo factible la realización de experimentos con muestras de números aleatorios en una computadora, siendo es aplicable tanto problemas estocásticos como determinísticos.

Para tener idea de su funcionamiento, consideremos un problema de probabilidad geométrica muy similar al planteado en el apartado de probabilidad clásica, sólo que ahora considerando regiones irregulares, donde la dificultad adicional consiste en calcular sus áreas.

Si la elección al azar de un punto del rectángulo se repite un gran número N de veces y se cuenta el número $N(R_i)$ de veces en que el punto cae en la región R_i, entonces el cociente $N(R_i)/N$ será una muy aceptable aproximación de la probabilidad $P(R_i)$ que mide la ocurrencia del evento R_i.

Si el área total del rectángulo base se denota con $A(R)$, entonces la expresión $A(R)N(R_i)/N$ proporciona una aproximación muy aceptable del área de la región R_i

Estas dos consideraciones constituyen la esencia del llamado método de Montecarlo, que en muchas aplicaciones concretas de ingeniería resulta operativamente más ventajoso que los métodos deterministas.

En 1733 Georges Louis Leclerc, Conde de Buffon, planteó dos problemas de probabilidad geométrica, con los que dejó evidencia de que los problemas de probabilidad también se podían resolver usando la geometría y no sólo la aritmética. Buffon ha sido reconocido como precursor de los llamados métodos de simulación de Montecarlo, que agrupan una serie de procedimientos de análisis probabilístico numérico.

El método de Montecarlo tomó su nombre del Casino del mismo nombre, del Principado de Mónaco, considerado capital mundial del juego, por ser la ruleta un generador natural de números aleatorios.

La invención del método de Montecarlo se les ha atribuido a Stanislaw Ulam y a John von Neumann, quienes en 1946 lo empezaron a usar como herramienta de investigación en el desarrollo de la bomba atómica, durante la Segunda Guerra Mundial; ellos simularon problemas probabilísticos de hidrodinámica relativos a la difusión de neutrones aleatorios en material de fusión.

A raíz del advenimiento y perfeccionamiento de la computadora, los métodos de simulación se han difundido mucho, especialmente por las velocidad de cálculo de las computadoras y por la calidad de las herramientas de soporte.

1.2.4 PROBABILIDAD SUBJETIVA

Cuando el fenómeno observado es intrínsecamente incierto, se requiere tomar decisiones guiándose por la previsión de eventos incontrolables que son los que generan tal incertidumbre, al menos en parte. La probabilidad aparece inmediatamente como respuesta, para medir esa incertidumbre y el modelo probabilístico que surge, abstrayendo la realidad, trae consigo una nueva incertidumbre, de otra índole, ya no intrínseca al fenómeno observado, sino asociada a la confianza en la información disponible sobre el fenómeno y a la calidad del modelo, en cuanto a representatividad de la realidad.

La previsión consiste en reconocer que existe incertidumbre, solo que ésta no es total, pues está acotada a un conjunto de resultados posibles, que implican riesgo, y que pueden ser ponderados de acuerdo a la evidencia empírica disponible o, cuando no la hay, conforme al grado de credibilidad racional subjetiva que se tenga, sobre su factibilidad. Para tomar decisiones, cabe poner especial atención en los resultados más factibles, que serían más desfavorables, a efecto de evitar que la decisión conlleve a efectos indeseables.

En el mundo real son mucho más frecuentes las situaciones en las que no es posible asignar probabilidades objetivas, basadas en evidencia, ya que los eventos, o son irrepetibles, o se pueden reproducir sólo unas cuantas veces.

La asignación de probabilidades a eventos no equiprobables, correspondientes a experimentos no repetibles, puede hacerse a través de lo que se conoce como interpretación subjetivista de probabilidad. La probabilidad de un evento puede ser simplemente una medida del grado de credibilidad que se tiene para la ocurrencia de un evento, expresado en términos numéricos.

Se ha desarrollado una serie de teorías subjetivistas, que consideran que la probabilidad de un evento es el grado de creencia que un sujeto tiene sobre la ocurrencia de ese evento, determinada a partir de su intuición, sentimiento, sentido común, experiencia o conocimiento.

Cuando una persona asigna probabilidad a uno de los posibles resultados de un proceso, con base en su conocimiento, experiencia o intuición, emite su propio juicio sobre la verosimilitud de que se obtenga ese resultado. Otra persona puede tener diferente opinión o información distinta y asignar una probabilidad diferente al mismo resultado. Se distinguen claramente dos escuelas de pensamiento en materia de probabilidad, relativas a grados de creencia.

Por un lado está la teoría de los grados de creencia racional y objetiva, conocida como interpretación logicista de probabilidad, que permite a un individuo estrictamente racional, asignar una medida de probabilidad a cualquier proposición lógica, expresando su grado de creencia en una hipótesis, sobre la base de una experiencia o una evidencia dada.

Una expresión relativa a probabilidades lógicas comparativas u ordinales se lee: la hipótesis H_1 condicionada por la evidencia E_1 no es más probable que la hipótesis H_2 condicionada por la experiencia E_2 y se expresa con la notación:

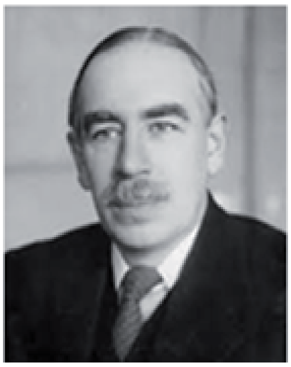
$$(H_1 \mid E_1) \prec (H_2 \mid E_2) \quad \Rightarrow \quad P(H_1 \mid E_1) \leq P(H_2 \mid E_2)$$

Una expresión relativa a una probabilidad lógica cuantitativa se lee: la función de confirmación C entre la hipótesis H y la evidencia E, expresa el grado de confirmación, con la notación: $C(H, E)$

La interpretación logicista de probabilidad no ofrece una base objetiva para que dos o más expertos que trabajan juntos obtengan una evaluación conjunta de su estado de conocimientos en un área científica de interés común. Si difieren los valores de las probabilidades asignadas por ellos, lo único que pueden hacer es someter a discusión sus argumentos e intentar llegar a un acuerdo.

Por otro lado está la teoría de los grados de creencia real, conocida como interpretación subjetivista de probabilidad, que permite a un individuo asignar una medida de probabilidad a cualquier proposición lógica, expresando su grado de creencia sobre la ocurrencia de un evento, con base en su intuición personal, sentimiento, sentido común, tal y como lo manifestaría en una apuesta de juego.

Una de las notaciones para la probabilidad subjetiva es $CS(H|E)$, que se lee: grado de creencia C que tiene el sujeto S, respecto a la hipótesis H, dada la evidencia E. El sujeto S asigna grados de creencia $CS(H|E)$ para distintos valores de H y E; las asignaciones incoherentes, si las hay, se eliminan, generando grados de creencia rectificados.



El iniciador de las teorías logicistas fue el economista Maynard Keynes, quien en 1921 publicó "*A Treatise on Probability*", que contiene el desarrollo de la sistematización del concepto de probabilidad lógica, como una conjetura inteligente, basada en el grado de creencia razonable establecido por la relación lógica entre dos proposiciones, mediante la asignación de probabilidad a la hipótesis, en función del soporte que ofrece la evidencia. Se requiere de un individuo idealmente racional que enjuicie la fuerza de la evidencia, desde la más absoluta imparcialidad, haciéndola equivaler a un determinado grado de probabilidad. Keynes consideró que, cuando es equilibrada la tendencia de los eventos a ocurrir o no ocurrir, estos son equiprobables, invocando el principio de indiferencia: "Todos los elementos de un conjunto discreto son indistinguibles cuando en ellos no se tiene en cuenta otra característica que el que los define como elementos de dicho conjunto; para establecer una distinción, hay que considerar otra característica no común a todos"

Los fundamentos de la teoría de probabilidad subjetiva fueron establecidos formalmente por Bruno De Finetti, quien en 1931 completó su desarrollo sistemático: "La probabilidad es una expresión de la opinión del observador del mundo y, como tal, no tiene existencia propia".

Sin embargo, fue hasta después de la publicación del libro de Leonard Savage, "*The Foundations of Statistics*", en 1954, cuando el concepto fue realmente aceptado y aplicado: "Para los personalistas la probabilidad mide la confianza que un cierto individuo tiene en la verdad de una proposición particular… y no niegan la posibilidad de que dos personas razonables, ante la misma evidencia, tengan grados de confianza distintos en la verdad de la misma proposición"

El enlace esencial para entender y aplicar la probabilidad subjetiva es la teoría de la utilidad, desarrollada por John von Neuman y Oskar Morgenstern. Lo que hace más fácil medir la probabilidad subjetiva que la probabilidad lógica, es la relación que aquella tiene con la utilidad y explica su gran aceptación.

Cuando se trabaja con probabilidad subjetiva, lo más importante es la consistencia de las asignaciones; el mismo sujeto, ante dos hipótesis lógicamente equivalentes, dadas las evidencias para ambas, lógicamente equivalentes también, asignará grados de creencia iguales; por lo general, la asignación subjetiva de probabilidades lleva consigo emociones, optimismo, pesimismo, deseo de que ocurra o de que no ocurra; todo eso se debe eliminar, de ser posible.

Las interpretaciones logicista y subjetivista presentan el inconveniente de que no parece humanamente posible que los juicios de una persona sobre las verosimilitudes relativas a un número infinito de sucesos sean completamente consistentes y libres de contradicciones.

Mientras que la probabilidad clásica y la probabilidad frecuencial de un evento se refieren a un valor fijo constante, la probabilidad subjetiva es una variable que puede y debe variar, en función de la nueva información recibida respecto del evento; esta manera de proceder, por cierto, es mucho más acorde a los lineamientos del método científico.

Ejemplo 1.21. BARCELONA - REAL MADRID. Para la asignación de probabilidades a los resultados de un partido de futbol, como ocurre para cualquier evento deportivo, difícilmente se puede suponer equiprobabilidad, por más equilibrados que estén los rivales en el momento del encuentro; obviamente, un partido de futbol no es como un volado, empezando por que el empate es factible, en tanto que es prácticamente imposible que la moneda caiga parada. En los 90 minutos que dura el encuentro se van a presentar varias secuencias de jugadas, en las que cada jugada podría considerarse semejante a un volado, en el sentido de que sale bien o mal para alguno de los dos equipos, por más que los directores técnicos intenten imponer una estrategia, hagan los cambios pertinentes o realicen ajustes en el acomodo de jugadores dentro de la cancha; dependiendo de este tipo de circunstancias, las secuencias pueden generar un partido plagado de emociones o del todo aburrido, un partido ganado mediante una sola jugada a balón parado, o con muchos goles producto de los errores del rival, logrados a base de bellas jugadas individuales, o un jugadas de conjunto, o con ausencia absoluta de goles. Ante un equilibrio tan patente, Laplace invocaría el principio de la razón insuficiente, considerando los tres resultados como igualmente posibles, pues no hay razón establecida para sospechar que un resultado del partido sea más probable que otro. Desde luego que este argumento no podría ser el mismo si el partido fuese España - Martinica.

Parece aún menos adecuado utilizar frecuencias relativas como aproximaciones de las probabilidades de los tres posibles resultados; por más veces que se hayan enfrentado en setenta años, o en el pasado más reciente, y por más que la estadística indique preponderancia a favor de uno de ellos, el partido próximo no tiene por qué parecerse a ninguno de los anteriores partidos y seguramente será diferente. Para este partido en particular, es posible que el análisis estadístico de resultados también indicara un equilibrio tal, que la asignación sería de $1/3$ para cada resultado y lo haría coincidente con el criterio clásico; queda claro que para este tipo de problemas, generalmente no es aplicable el criterio frecuentista.

Finalmente, queda la asignación de probabilidades subjetivas, que son sin duda las más socorridas en estos casos; ésta la puede hacer desde el analista más reconocido, enterado de todo lo que sucede en torno al futbol, hasta la señora que vende quesos en el mercado y que no tiene ni idea de futbol y menos aún de los equipos Barcelona y Real Madrid; quizá haya que explicarle que un 100% habrá que distribuirlo en tres resultados; gana Real Madrid, gana Barcelona y empatan, y una vez asignadas las probabilidades, verificar que éstas sean consistentes. El analista experto, en cambio, seguramente tomará en cuenta que no podrán alinear algunos jugadores lesionados o castigados, que tal o cual sistema de juego le puede funcionar mejor a cada equipo, las posibles necedades de los técnicos, la cancha en la que se juega, la hora del partido, las posibilidades de lluvia, etc., y luego, conforme a su experiencia y conocimientos, asignará las probabilidades correspondientes. Ahora que lo más probable es, que siendo conocedor de futbol, también sea fanático de alguno de los dos equipos, y entonces, su asignación llevará consigo una dosis de deseo, que seguramente sesgará su pronóstico en favor de su equipo favorito. Paradójicamente, no sería nada raro que de todas las asignaciones realizadas con distintos criterios, la más atinada sea la de la señora de los quesos.

1.2.5 TEOREMAS FUNDAMENTALES

La probabilidad tuvo su propia evolución histórica, muy singular, comparativamente con la de otras ramas de la matemática. Su conceptualización formal fue la última de todas, apenas en 1933, quizá porque su aceptación como disciplina matemática fue tardía. Sin embargo, no es creíble que apenas hace tres siglos y medio, el hombre haya empezado a abordar el azar desde la perspectiva matemática. Si bien con Pascal y Fermat se inició en Occidente su desarrollo acelerado, impulsado en primera instancia por los jugadores y luego por los demógrafos y las compañías de seguros, es de elemental justicia reconocer las aportaciones de babilonios, egipcios, chinos y romanos, así como los grandes logros que, durante los 1000 años de oscurantismo europeo, se alcanzaron en la India y luego con los árabes, en concepciones intuitivas, desarrollos teóricos, terminología y cálculos de probabilidades.

Al igual que todas las ciencias, la probabilidad también atravesó por un período empírico de consolidación, forjado desde la perspectiva frecuentista, que sigue demandando recurrencia a la observación, porque la probabilidad tiene una componente inferencial preponderante; su desarrollo teórico deductivo no fue posterior, sino que se inició casi en forma simultánea, atendiendo el punto de vista del jugador, aparentemente alejado de la realidad, con el supuesto de resultados igualmente posibles.

Ante la necesidad de acabar con la especulación sobre si la población aumentaba, disminuía o permanecía estática, los precursores de la estadística introdujeron el concepto de frecuencia de un evento y encontraron, por ejemplo, que regularmente nacían más hombres que mujeres y que había una clara variación estacional en la ocurrencia de las muertes, dándose cuenta, además, que cuantas más observaciones hacían, más exactos y precisos eran sus resultados, anticipando el principio de estabilidad de las frecuencias.

Las concepciones logicista y subjetivista han sido utilizadas, de hecho, desde tiempos inmemoriales, si bien sus definiciones formales han sido recientes. Se considera que estas dos concepciones y la frecuentista son virtualmente incompatibles; sin embargo, han sido estudiadas aplicadas y valoradas por muchos investigadores y la utilidad de cada una de ellas ha sido reconocida.

También hubo otros que se percataron que al suponer que los sucesos reales tienen una posibilidad de ocurrencia, que puede ser medida en forma aproximada, se podía deducir un conjunto de premisas que eran aplicables también aproximadamente a esos sucesos reales. Ese conjunto de proposiciones lógicas, denominado teoría de la probabilidad, modelada sobre la realidad, pero independiente de esa realidad, dio la posibilidad de deducir nuevas proposiciones, que experimentalmente sólo habrían sido encontradas por casualidad.

Todas las ramas de la matemática parten de un pequeño conjunto de definiciones y axiomas, que se transforma en un conjunto grande de proposiciones, denominados teoremas. Aunque pareciera que cada una de las proposiciones obtenidas está implícita en los axiomas, en la realización de cada transformación está presente un proceso creativo.

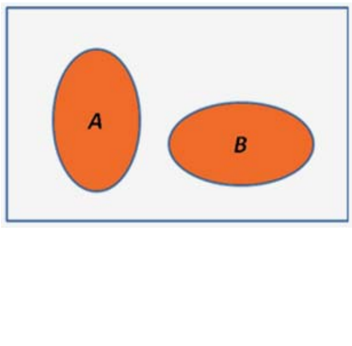
Toda la teoría de probabilidad se puede construir, por ejemplo, a partir de los siguientes tres principios básicos, que podrían ser llamados los teoremas fundamentales de la probabilidad, los cuales se enuncian y se demuestran de la manera muy simple.

Teorema de recorrido

No obstante que las tres interpretaciones discutidas son distintas en forma y en fondo, la probabilidad de un evento A puede verse como el resultado de un cociente de números no negativos, cuyo numerador es al menos cero y a lo sumo igual al denominador; entonces, la probabilidad de un evento es una función cuyo recorrido fluctúa entre cero y uno: $0 \leq P(A) \leq 1$

Esta simple condición constituye un teorema que, aunque no lo parezca, ha quedado enunciado y demostrado en el párrafo anterior. Si el evento es imposible, el numerador es igual a cero y la probabilidad es nula; si el evento es seguro, el numerador es igual al denominador y la probabilidad es uno.

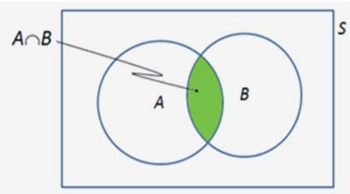
Teorema de probabilidad total



Si al realizar un experimento, un evento C puede ocurrir en las modalidades incompatibles A y B, entonces la probabilidad de ocurrencia de C es la suma de las respectivas probabilidades de A y B:

$$A \cap B = \phi \quad \Rightarrow \quad P(C) = P(A \cup B) = P(A) + P(B)$$

La incompatibilidad entre los eventos es muy fácil de advertir, porque al excluirse entre sí, no pueden ocurrir conjuntamente. Evidentemente, cuando los eventos no son mutuamente exclusivos, el enunciado cambia y se requiere usar otro teorema, que considera la ocurrencia conjunta.



Teorema de probabilidad conjunta

Si la ocurrencia del evento A no influye en la probabilidad de ocurrencia del evento B, y viceversa, se dice que los eventos A y B son independientes y, entonces, la probabilidad de que los dos eventos ocurran de manera conjunta es igual al producto de sus respectivas probabilidades:

$$A \text{ y } B \text{ independientes} \quad \Rightarrow \quad P(A \cap B) = P(A)P(B), \quad P(A) \neq 0 \neq P(B)$$

La independencia entre eventos no es una propiedad que pueda identificarse fácilmente; si los eventos son físicamente independientes, entonces también son estadísticamente independientes, pero suele suceder que la independencia física es solo aparente. Evidentemente, cuando los eventos no son independientes, el enunciado cambia y se tiene que usar otro teorema, que considere la dependencia entre los eventos.

Ninguno de los tres teoremas fundamentales de la probabilidad requieren de altas matemáticas; usan únicamente suma y producto de dos números.

Lo delicado de estos teoremas está en su correcta aplicación, lo que implica distinguir eventos incompatibles y compatibles, por un lado, y eventos independientes y dependientes, por otro.

Ejemplo 1.22. BARAJA INGLESA. Si de una baraja inglesa de 52 cartas se extraen sucesivamente dos cartas al azar, haciendo cada vez el reemplazo de las cartas extraídas, calcule la probabilidad de que en la primera extracción la carta sea un as y en la segunda extracción la carta sea un corazón.

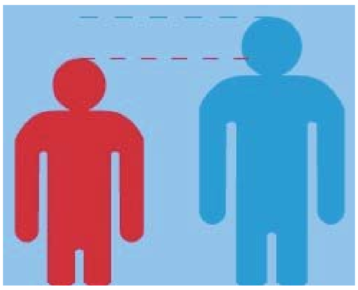
Se está pidiendo la probabilidad de ocurrencia conjunta para la obtención de un as y luego de un corazón, en la extracción sucesiva de dos cartas. Puesto que hay reemplazo, los eventos son independientes, porque en cada una de las extracciones hay 52 cartas disponibles, con 4 ases y 13 corazones.

Por tratarse de un juego de azar, procede aplicar el criterio de Laplace, considerando los eventos: A = {primera carta es un as}; $P(A) = 4/52 = 1/13$

C = {segunda carta es corazón}; $P(C) = 13/52 = 1/4$

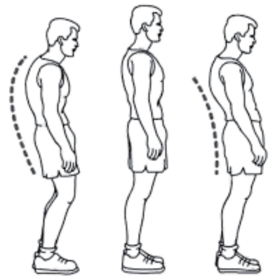
La probabilidad conjunta es: $P(A \cap C) = P(A)P(C) = \dfrac{1}{13} \cdot \dfrac{1}{4} = \dfrac{1}{52}$

Si no hubiera reemplazo, los eventos ya no serían independientes, porque el hecho de que la primera carta sea un as, afecta la probabilidad de que la segunda sea un corazón, pues en ésta ya solo hay 51 cartas disponibles y el número de corazones puede seguir siendo 13 o se pudo haber reducido a 12, si el as que salió fue el de corazones. Analizaremos esta circunstancia y su solución, en el capítulo 1.5.



Ejemplo 1.23. ESTATURA.

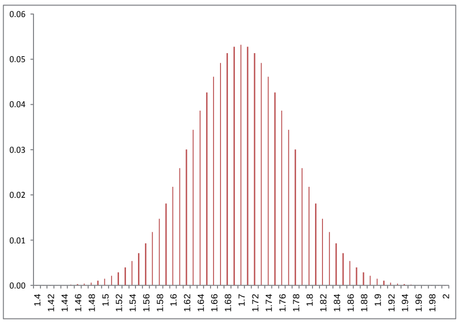
La estatura humana es una variable que depende de la genética, el medio ambiente, la edad y el sexo; la estatura media es una característica típica en un grupo que forma parte de una población con antecedentes genéticos y factores ambientales comunes. En México, por ejemplo, la estatura media de los adultos varones es de 1.70 m, en tanto que en el centro y en el sur, el promedio baja a 1.60 m.

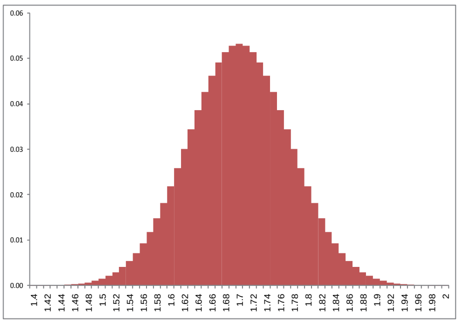


Cada vez que se toma la estatura de una persona se pueden estar cometiendo tres tipos de errores: el primero corresponde a las variaciones que tiene la estatura, por la temperatura ambiental y por la posición que adopta la persona al ser medida; el segundo es debido a la precisión que puede ofrecer el instrumento de medición; y el tercero es el que se comete por el redondeo.

Aunque no sea cierto, la estatura puede ser manejada como una variable estrictamente discreta, dejando establecido que será expresada con tres cifras significativas, dos de ellas decimales, y no podrá haber estaturas con tres o más decimales.

Por supuesto que la estatura es una variable continua; cuando tú dices que mides 1.72 m, estás diciendo que tu estatura está realmente entre 1.715 y 1.725 m, y cuando yo digo que mido 1.79 m, estoy dando una aproximación de mi real estatura, que está entre 1.785 y 1.795 m. Al manejar la estatura como variable continua, expresándola en forma discretizada, con aproximación de dos decimales, cada valor representa todos los valores de un intervalo de un centímetro de longitud; y así se manejan todas las variables continuas con las que tratamos los ingenieros: velocidad de un vehículo, voltaje en una toma de corriente, volumen de líquido contenido en un tanque, etc.; siempre que medimos, nos vemos obligados a discretizar. Ahora bien, parece muy sutil la diferencia en la forma de representar gráficamente las frecuencias relativas considerando la estatura como variable discreta y como variable continua; en el segundo capítulo veremos que existen diferencias fundamentales.





En cualquiera de los dos casos, la probabilidad de que un individuo elegido al azar mida 1.72 m es de 0.0513; la probabilidad de que un segundo individuo elegido al azar mida 1.72 m es de 0.0513. Obviamente, si los eventos son independientes, la probabilidad de que ambos midan 1.72 m es el producto de las probabilidades individuales: $0.0513 \times 0.0513 = 0.00263$. Pero qué ocurre si esos dos individuos son de la misma raza, o del mismo pueblo, o de la misma familia, o si son hermanos, o hasta si son gemelos. Se puede percibir claramente que, a medida que el vínculo entre ellos es mayor, parece natural que sus estaturas se asemejen cada vez más y, por lo tanto, los eventos se vayan haciendo cada vez más dependientes, hasta el punto de que la probabilidad conjunta de que ambos midan 1.72 m se va acercando a $0.0513 \times 1 = 0.0513$.

1.2.6 PROBLEMAS CLÁSICOS

Martingala

En la ruleta, apostando solamente al rojo o al negro, si aciertas el color, ganas el doble de tu apuesta, y si fallas, simplemente la pierde. Algunos jugadores creen que hay un método infalible para ganar, conocido como *martingala*, el cual consiste en doblar la apuesta cada vez que pierdes. Si tu apuesta inicial es de 1 euro, si ganas guardas tus 2 euros y te vas, y si pierdes, tu siguiente apuesta será de 2 euros, si ganas guardas tus 4 euros y te vas, y si pierdes, tu siguiente apuesta será de 4 euros, y sólo si pierdes, seguirás apostando el doble de la anterior: 8 euros, 16 euros, etc., hasta que ganes una. Según los seguidores del método, a la larga, siempre ganarás 1 euro, tu apuesta inicial.

Lo dicho parece cierto siempre, porque si pierdes la primera (-1) y ganas la segunda (+2), tu ganancia es de 1 euro; si pierdes las dos primeras (-3) y ganas la tercera (+4), tu ganancia es de 1 euro, si pierdes las tres primeras (-7) y ganas la cuarta (+8), tu ganancia es de 1 euro; y así sucesivamente; el problema está cuando pierdes las cinco primeras (-31) y en tu bolsa sólo había 31 euros; pues los pierdes. El quid del asunto está en que tú solo dispones de un número limitado de euros y por grande que sea tu capital, obviamente es finito.

Se puede argumentar que es muy poco probable perder 5 veces seguidas, y es cierto, pues si en cada tirada tu probabilidad de ganar o perder es 1/2, en 5 tiradas tu probabilidad de ganar es de 31/32 y la de perder es de 1/32; si acertaras en esa quinta tirada, tu ganancia acumulada sería de 1 euro, y si fallaras, tu pérdida acumulada sería de -31 euros. Se puede observar claramente el equilibrio: $(31/32)(1) = 31/32$; $(1/32)(-31) = -31/32$; el juego parece ser completamente equitativo, lo cual haría que, jugando a la ruleta con este método, a la larga, ni ganas ni pierdes.

Los cálculos anteriores producen los mismos resultados, independientemente de cuál haya sido tu apuesta inicial; si la aumentas, lo único que harás es reducir el número de apuestas que puedes soportar con tu disponible, sin que por ello se modifique pérdidas o ganancias; y si la reduces, tampoco habrá cambio en el resultado y sólo te quedarás más tiempo en la mesa de juego.

Suponiendo que todos los jugadores utilizaran este método para jugar, unos apostarían al rojo y otros al negro y, según parece ninguno ganaría, ni perdería. Pero recordemos que la banca siempre gana, así que el juego no puede ser equitativo, aunque lo parezca; la ruleta tiene 18 ranuras para el rojo y 18 ranuras para el negro, pero hay 1 ranura adicional para el verde, que es el cero; así que tu probabilidad de ganar en cada tirada no es 1/2 sino 18/37; y si estás en un casino de Estados Unidos, son 2 las ranuras verdes: una para el cero y otra para el doble cero, y tu probabilidad se reduce a 9/19. Esas son la ventajas de entrada que tienen las respectivas bancas, respectivamente, en Europa y en Estados Unidos: 2.7% y 5.26% de probabilidades de ganar, más que tú y que cualquier otro jugador.

Y la banca tiene otra ventaja y es que el casino te pone un tope para hacer una apuesta, de modo que aún teniendo mucho dinero para seguir doblando tu apuesta, no podrías hacerlo y terminarías perdiendo todo lo apostado.

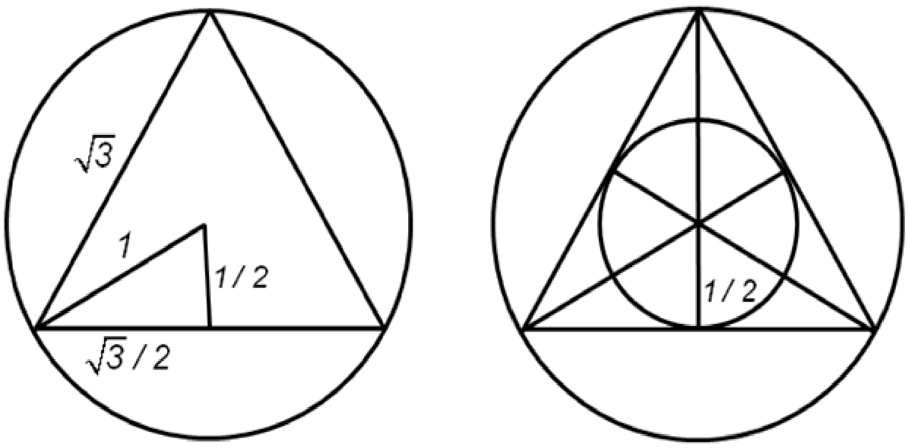
Por extensión se le llama martingala a todo medio seguro de ganar; los matemáticos han demostrado que las martingalas son como las brujas, o los dragones; no existen.

Paradoja de Bertrand

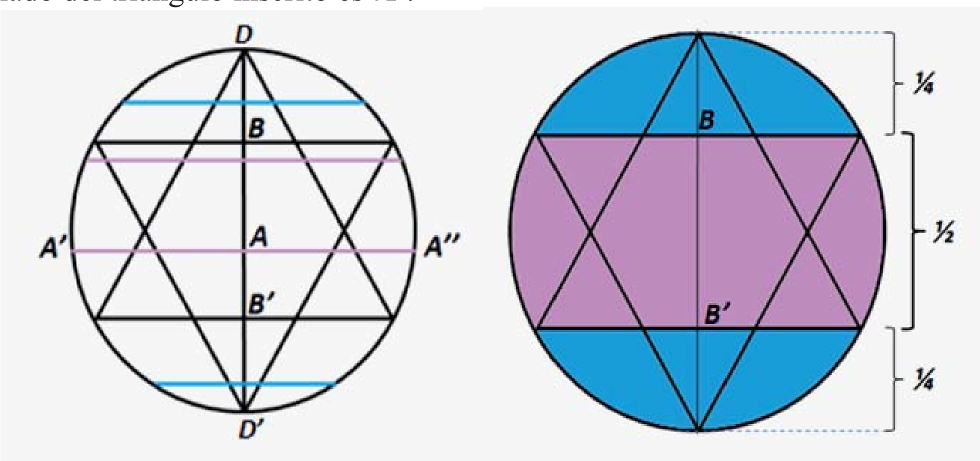
Considere un triángulo equilátero inscrito en un círculo y suponga que se traza al azar una cuerda del círculo. Determine la probabilidad de que esa cuerda sea más larga que un lado del triángulo. Resuelva el problema considerando cada una de las siguientes tres hipótesis:

1. Segmentos con la misma longitud tienen la misma probabilidad de ocurrir.
2. Sectores con el mismo ángulo tienen la misma probabilidad de ocurrir.
3. Regiones con la misma área tienen la misma probabilidad de ocurrir.

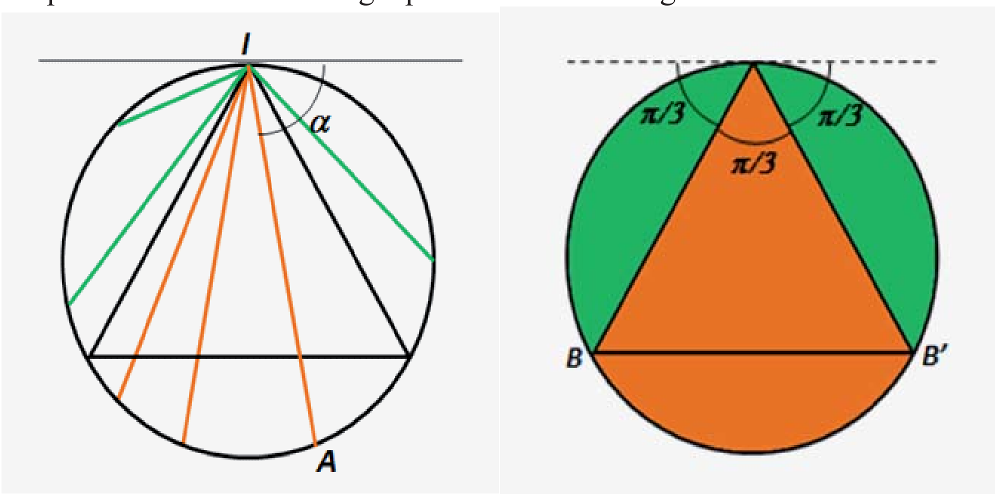
Con recursos de geometría y trigonometría, se determina fácilmente que el lado del triángulo equilátero inscrito en una circunferencia de radio *1* es y el radio de la circunferencia inscrita en el triángulo inscrito es *1/2*.



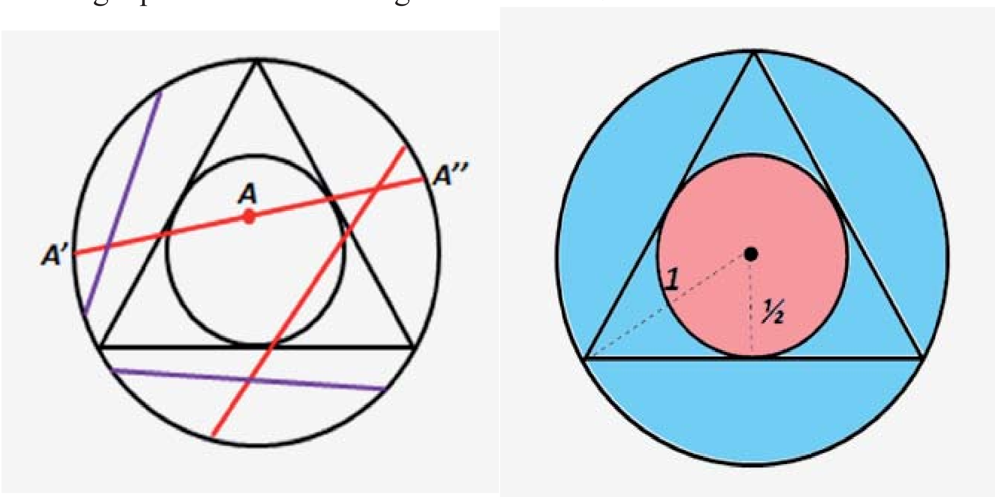
1. Se escoge un diámetro *DD*' cualquiera del círculo, luego se elige al azar un punto *A* sobre ese diámetro y, perpendicular a éste, se traza una cuerda *A'A''* que pasa por el punto *A*. Para facilitar los cálculos, se sobrepone otro triángulo rotado, de manera que uno de sus lados quede perpendicular al diámetro seleccionado. La cuerda será más larga que un lado del triángulo sólo si se escoge un punto cercano al centro del círculo, dentro del intervalo *BB'* delimitado por las intersecciones del diámetro con los lados de los dos triángulos, cuya longitud es ½ . Por lo tanto la probabilidad de que la cuerda sea más larga que un lado del triángulo inscrito es ½ .



2. Se selecciona el vértice *I* del triángulo inscrito, se escoge al azar un punto *A* de la circunferencia y se traza la cuerda *IA* que une esos dos puntos, formando un ángulo α con la recta tangente a la circunferencia, en el punto *I*. La cuerda será más larga que el lado del triángulo, solo si se escoge un punto *A* que pertenezca al arco delimitado por los puntos BB’ que definen el lado opuesto del triángulo, arco cuya ángulo es un tercio de π. Por lo tanto, la probabilidad de que la cuerda sea más larga que un lado del triángulo inscrito es 1/3.



3. Se escoge un punto A cualquier del círculo y se traza la cuerda *A’A*’’ cuyo punto medio es el punto seleccionado. La cuerda es más larga que un lado del triángulo inscrito si el punto A está en el círculo inscrito en el triángulo inscrito, y que tiene por radio ½; el área de ese pequeño círculo tiene la cuarta parte del área del círculo unitario. Por lo tanto, la probabilidad de que la cuerda sea más larga que un lado del triángulo inscrito es ¼.



Hemos encontrado tres soluciones diferentes para un mismo problema y ninguna de ellas es mejor que las otras; y se pueden construir otras soluciones, basadas en otras maneras de elegir una cuerda al azar. Los resultados obtenidos son distintos porque las hipótesis de partida son distintas; conviene que revisemos en cada caso el procedimiento seguido para elegir una cuerda al azar.

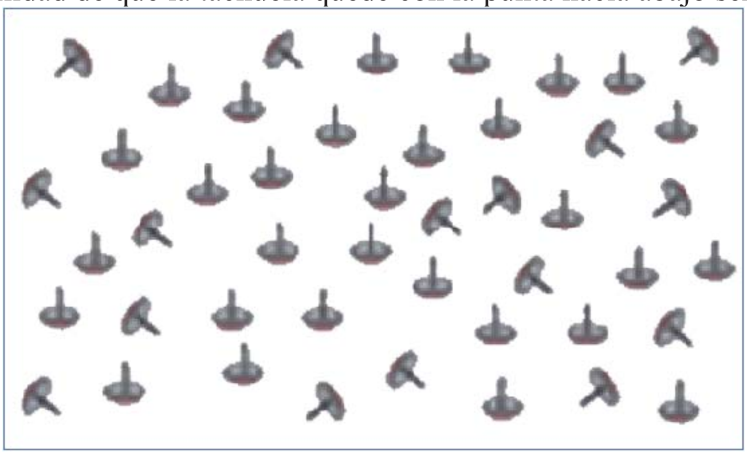
En todo problema de probabilidad geométrica, se invoca el criterio de Laplace, que exige simetría de los resultados elementales; pero esta conceptualización a priori se puede hacer extensiva a espacios infinitos, sólo si los conjuntos elementales tienen la misma medida geométrica, para poder considerarlos como equiprobables. En la primera hipótesis se habla de segmentos de igual longitud, en la segunda se hace referencia a sectores con el mismo ángulo y en la tercera se consideran regiones con la misma área. La equiprobabilidad en longitud, ángulo y área nos conducen a diferentes procedimientos y a diferentes resultados.

Lanzamiento de tachuelas

Si lanzamos una tachuela al aire, cuando cae al suelo y se queda quieta, hay dos posibilidades: la punta está hacia arriba o la punta está hacia abajo. Los resultados no parecen equiprobables, de modo que la interpretación clásica, con una asignación 50-50% no es aplicable. Podemos recurrir entonces a la frecuencia relativa o a la probabilidad subjetiva.

Sin hacer experimento alguno y tomando en cuenta únicamente las característica observables de la tachuela, cada uno de los presentes asigne individualmente probabilidades a los eventos, a una cifra decimal; luego de discutirlo con su compañero de al lado, acuerden entre ambos la asignación de probabilidades con dos decimales, a lo más. Una vez vaciadas estas probabilidades estimadas en el pizarrón, intentemos llegar a un consenso.

Es posible realizar el experimento repetidamente en condiciones similares y de manera independiente. Considerando el experimento consistente en lanzar 10 tachuelas simultáneamente, si después de haberlo llevado a cabo 40 veces, y habiendo contado cada vez, la frecuencia de cada posición, resultó que de las 400 tachuelas lanzadas, hubo en total 170 que quedaron con la punta hacia arriba, lo cual corresponde a una frecuencia relativa de 0.425, que puede ser considerada como una buena aproximación de la probabilidad; obviamente, la probabilidad de que la tachuela quede con la punta hacia abajo sería 0.575.



Vemos que las asignaciones subjetivistas y frecuentistas no necesariamente se parecen, y si así fuera, sería más producto de la casualidad que de la intuición colectiva. Obviamente, siempre se preferirá la frecuencia relativa, cuando sea factible obtenerla.

Sin embargo, la evidencia empírica obtenida no nos permite declarar, así como así, que la tachuelas caen con la punta hacia arriba en el 42.5% de los casos y hacia abajo el 57.5%, simplemente porque existen diferentes tipos de tachuelas, con muchos clases de cabezas en forma y en tamaño, así como distintas longitudes de puntas.

Tomando en cuenta la evidencia empírica obtenida previamente, asignemos ahora probabilidades subjetivas al evento "punta hacia arriba", para los cuatro tipos de tachuelas mostrados en la fotografía.

Prototipo

A una compañía de diseño industrial se le ha encargado la construcción de un cierto prototipo, que de resultar satisfactorio, representaría para la empresa un sustancioso beneficio, pero en caso contrario implicaría una pérdida. El gerente desea saber cuál es la probabilidad de que el prototipo que se diseñe sea aprobado por el cliente, a fin de decidir si le conviene a la compañía aceptar la propuesta. En este caso no se cuenta con información estadística, pues jamás se había hecho algo similar; entonces, ¿cómo asignarle probabilidad al evento Se tiene que recurrir al experto, al ingeniero mecánico en este caso, el que en base en su intuición y experiencia dirá, por ejemplo: "la probabilidad de que el prototipo resulte satisfactorio es 0.8.

Suponga que son dos los expertos que pueden opinar sobre el diseño del prototipo y que habiendo asignado originalmente una probabilidad de éxito de 0.8, el primero, y 0.5, el segundo, después de discutirlo, llegaron a la conclusión de que tal probabilidad estaba entre 0.6 y 0.7.

En ocasiones es necesario afinar la subjetividad del experto a través de los conceptos de preferencia e indiferencia de la teoría de la utilidad. La técnica consiste en someter a la consideración del experto la elección de uno de dos juegos. El primer juego consiste en que, si el prototipo es aceptado, el experto gana y si el prototipo es rechazado, el experto pierde. El segundo juego consiste en extraer una bola de una urna que tiene un total de 100 bolas, x de las cuales son blancas y $(100 – x)$ son negras; si la bola extraída es blanca, el experto gana, y si es negra, pierde. Para ambos juegos, si el experto gana obtiene un premio de 30 días en Europa, por ejemplo; y si el experto pierde, obtiene un castigo de 30 días en el CERESO de Gómez Palacio, por ejemplo.

Suponga que el gerente le propone tal elección al experto de su confianza, con una urna en la que hay 70 bolas blancas y 30 negras; suponga que él prefiere el segundo juego, lo cual significa que él siente que tiene mayor oportunidad de ganar con la urna, que con el prototipo. Entonces el gerente le cambia la urna, ahora con 60 bolas blancas y 40 negras, y ahora el experto prefiere el primer juego, pues le parece más fácil ganar con el prototipo que con la urna.

El gerente le va proponiendo diferentes urnas: con 69 bolas blancas y 31 negras, con 61 bolas blancas y 39 negras, con 68 bolas blancas y 32 negras, y así sucesivamente, hasta que el experto se muestra indiferente entre los dos juegos; suponga que esto ocurre cuando la urna tiene 64 bolas blancas y 36 negras; entonces, la probabilidad subjetiva del experto de que el prototipo resulte satisfactorio es 0.64.

La segunda parte de la técnica consiste en corroborar el resultado, tratando que el experto rehuya el castigo, y no como lo había hecho antes, buscando el premio. Ahora en el primer juego, se otorga premio si el evento de interés no ocurre y castigo si ocurre; y en el segundo juego, se otorga premio si la bola extraída es negra, y castigo, si es blanca. De obtenerse el mismo resultado que en la primera parte, se tendría una medida de probabilidad, que aunque intuitiva, es bastante confiable. De obtenerse otro resultado, habrá que repetir el ejercicio hasta obtener el ajuste deseado.

Puede verse que la interpretación subjetiva de la probabilidad puede ser formalizada, cuando los juicios de una persona acerca de las verosimilitudes relativas a diversas combinaciones de resultados satisfacen ciertas condiciones de consistencia; entonces, sus probabilidades subjetivas para los diferentes sucesos posibles pueden ser determinadas en forma única.