Probabilidades Doctas con Discos, Árboles, Bolitas y Urnas

ISBN: 978-956-306-071-3
Registro de Propiedad Intelectual: 200.531
Colección: Herramientas para la formación de profesores de matemáticas.
Diseño: Jessica Jure de la Cerda
Diseño de Ilustraciones: Catalina Frávega Thomas, Cristina Felmer Plominsky
Diagramación: Pedro Montealegre Barba, Francisco Santibáñez Palma
Financiamiento: Proyecto Fondef D05I-10211

Datos de contacto para la adquisición de los libros:
Para Chile:
1. En librerías para clientes directos.
2. Instituciones privadas directamente con:
   Juan Carlos Sáez C.
   Director Gerente
   Comunicaciones Noreste Ltda.
   J.C. Sáez Editor
   jcsaezc@vtr.net
   www.jcsaezeditor.blogspot.com
   Oficina: (56 2) 3260104 - (56 2) 3253148
3. Instituciones públicas o fiscales: www.chilecompra.cl

Desde el extranjero:
1. Liberalia Ediciones: www.liberalia.cl
2. Librería Antártica: www.antartica.cl
3. Argentina: Ediciones Manantial: www.emanantial.com.ar
4. Colombia: Editorial Siglo del Hombre
   Fono: (571) 3377700
5. España: Tarahumara, tarahumara@tarahumaralibros.com
   Fono: (34 91) 3656221
6. México: Alejandría Distribución Bibliográfica, alejandria@alejandrialibros.com.mx
   Fono: (52 5) 556161319 - (52 5) 6167509
7. Perú: Librería La Familia, Avenida República de Chile # 661
8. Uruguay: Dolmen Ediciones del Uruguay
   Fono: 00-598-2-7124857

Probabilidades Doctas: con Discos, Árboles y Urnas — Pierre Paul Romagnoli
Universidad Andrés Bello
promagnoli@unab.cl

ESTA PRIMERA EDICIÓN DE 2.000 EJEMPLARES
Se terminó de imprimir en febrero de 2011 en WORLDCOLOR CHILE S.A.

PROBABILIDADES DOCTAS CON DISCOS, ÁRBOLES, BOLITAS Y URNAS

Pierre Paul Romagnoli

Universidad Andrés Bello

Dirección del Proyecto Fondef D05I-10211
Herramientas para la Formación de Profesores de Matemática

Patricio Felmer, Director del Proyecto
Universidad de Chile.

Leonor Varas, Directora Adjunta del Proyecto
Universidad de Chile.

Salomé Martínez, Subdirectora de Monografías
Universidad de Chile.

Cristián Reyes, Subdirector de Estudio de Casos
Universidad de Chile.

Presentación de la Colección

La colección de monografías que presentamos es el resultado del generoso esfuerzo de los autores, quienes han dedicado su tiempo y conocimiento a la tarea de escribir un texto de matemática. Pero este esfuerzo y generosidad no se encuentra plenamente representado en esta labor, sino que también en la enorme capacidad de aprendizaje que debieron mostrar, para entender y comprender las motivaciones y necesidades de los lectores: Futuros profesores de matemática.

Los autores, encantados una y otra vez por la matemática, sus abstracciones y aplicaciones, enfrentaron la tarea de buscar la mejor manera de traspasar ese encanto a un futuro profesor de matemática. Éste también se encanta y vibra con la matemática, pero además se apasiona con la posibilidad de explicarla, enseñarla y entregarla a los jóvenes estudiantes secundarios. Si la tarea parecía fácil en un comienzo, esta segunda dimensión puso al autor, matemático de profesión, un tremendo desafío. Tuvo que salir de su oficina a escuchar a los estudiantes de pedagogía, a los profesores, a los formadores de profesores y a sus pares. Tuvo que recibir críticas, someterse a la opinión de otros y reescribir una y otra vez su texto. Capítulos enteros resultaban inadecuados, el orden de los contenidos y de los ejemplos era inapropiado, se hacía necesario escribir una nueva versión y otra más. Conversaron con otros autores, escucharon sus opiniones, sostuvieron reuniones con los editores. Escuchar a los estudiantes de pedagogía significó, en muchos casos, realizar eventos de acercamiento, desarrollar cursos en base a la monografía, o formar parte de cursos ya establecidos. Es así que estas monografías recogen la experiencia de los autores y del equipo del proyecto, y también de formadores de profesores y estudiantes de pedagogía. Ellas son el fruto de un esfuerzo consciente y deliberado de acercamiento, de apertura de caminos, de despliegue de puentes entre mundos, muchas veces, separados por falta de comunicación y cuya unión es vital para el progreso de nuestra educación.

La colección de monografías que presentamos comprende una porción importante de los temas que usualmente encontramos en los currículos de formación de profesores de matemática de enseñanza media, pero en ningún caso pretende ser exhaustiva. Del mismo modo, se incorporan temas que sugieren nuevas formas de abordar los contenidos, con énfasis en una matemática más pertinente para el futuro profesor, la que difiere en su enfoque de la matemática para un ingeniero o para un licenciado en matemática, por ejemplo. El formato de monografía, que aborda temas específicos con extensión moderada, les da flexibilidad para que sean usadas de muy diversas maneras, ya sea como texto de un curso, material complementario, documento básico de un seminario, tema de memoria y también como lectura personal. Su utilidad ciertamente va más allá de las aulas universitarias, pues esta colección puede convertirse en la base de una biblioteca personal del futuro profesor o profesora, puede ser usada como material de consulta por profesores en ejercicio y como texto en cursos de especialización y post-títulos. Esta colección de monografías puede ser usada en concepciones curriculares muy distintas. Es, en suma, una herramienta nueva y valiosa, que a partir de ahora estará a disposición de estudiantes de pedagogía en matemática, formadores de profesores y profesores en ejercicio.

El momento en que esta colección de monografías fue concebida, hace cuatro años, no es casual. Nuestro interés por la creación de herramientas que contribuyan a la formación de profesores de matemática coincide con un acercamiento entre matemáticos y formadores de profesores que ha estado ocurriendo en Chile y en otros lugares del mundo. Nuestra motivación nace a partir de una creciente preocupación en todos los niveles de la sociedad, que ha ido abriendo paso a una demanda social y a un interés nacional por la calidad de la educación, expresada de muy diversas formas. Esta preocupación y nuestro interés encontró eco inmediato en un grupo de matemáticos, inicialmente de la Universidad de Chile, pero que muy rápidamente fue involucrando a matemáticos de la Pontificia Universidad Católica de Chile, de la Universidad de Concepción, de la Universidad Andrés Bello, de la Universidad Federico Santa María, de la Universidad Adolfo Ibáñez, de la Universidad de La Serena y también de la Universidad de la República de Uruguay y de la Universidad de Colorado de Estados Unidos.

La matemática ha adquirido un rol central en la sociedad actual, siendo un pilar fundamental que sustenta el desarrollo en sus diversas expresiones. Constituye el cimiento creciente de todas las disciplinas científicas, de sus aplicaciones en la tecnología y es clave en las habilidades básicas para la vida. Es así que la matemática actualmente se encuentra en el corazón del currículo escolar en el mundo y en particular en Chile. No es posible que un país que pretenda lograr un desarrollo que involucre a toda la sociedad, descuide el cultivo de la matemática o la formación de quienes tienen la misión de traspasar de generación en generación los conocimientos que la sociedad ha acumulado a lo largo de su historia.

Nuestro país vive cambios importantes en educación. Se ha llegado a la convicción que la formación de profesores es la base que nos permitirá generar los cambios cualitativos en calidad que nuestra sociedad ha impuesto. Conscientes de que la tarea formativa de los profesores de matemática y de las futuras generaciones de jóvenes es extremadamente compleja, debido a que confluyen un sinnúmero de factores y disciplinas, a través de esta colección de monografías, sus editores, autores y todos los que han participado del proyecto en cada una de sus etapas, contribuyen a esta tarea, poniendo a disposición una herramienta adicional que ahora debe tomar vida propia en los formadores, estudiantes, futuros profesores y jóvenes de nuestro país.

Patricio Felmer y Salomé Martínez
Editores

Agradecimientos

Agradecemos a todos quienes han hecho posible la realización de este proyecto Fondef: "Herramientas para la formación de Profesores de Matemáticas". A Cristián Cox, quien apoyó con decisión la idea original y contribuyó de manera crucial para obtener la participación del Ministerio de Educación como institución asociada. Agradecemos a Carlos Eugenio Beca por su apoyo durante toda la realización del proyecto. A Rafael Correa, Edgar Kausel y Juan Carlos Sáez, miembros del Comité Directivo. Agradecemos a Rafael Benguria, Servet Martínez y Fidel Oteiza, miembros del Comité Editorial de la colección, quienes realizaron valiosos aportes a los textos. A Guillermo Marshall, Decano de la Facultad de Matemáticas de la Pontificia Universidad Católica de Chile y José Sánchez, entonces Decano de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Concepción, quienes contribuyeron de manera decisiva a lograr la integridad de la colección de 15 monografías. A Jaime San Martín, director del Centro de Modelamiento Matemático por su apoyo durante toda la realización del proyecto. Agradecemos a Víctor Campos, Ejecutivo de Proyectos de Fondef, por su colaboración y ayuda en las distintas etapas del proyecto.

Agradecemos también a Bárbara Ossandón de la Universidad de Santiago, a Jorge Ávila de la Universidad Católica Silva Henríquez, a Víctor Díaz de la Universidad de Magallanes, a Patricio Canelo de la Universidad de Playa Ancha en San Felipe y a Osvaldo Venegas y Silvia Vidal de la Universidad Católica de Temuco, quienes hicieron posible las visitas que realizamos a las carreras de pedagogía en matemática. Agradecemos a todos los evaluadores, alumnos, académicos y profesores -cuyos nombres no incluimos por ser más de una centena- quienes entregaron sugerencias, críticas y comentarios a los autores, que ayudaron a enriquecer cada uno de los textos.

Agradecemos a Marcela Lizana por su impecable aporte en todas las labores administrativas del proyecto, a Aldo Muzio por su colaboración en la etapa de evaluación, y también a Anyel Alfaro por sus contribuciones en la etapa final del proyecto y en la difusión de los logros alcanzados.

Dirección del Proyecto

Índice General

Prefacio

Esta monografía está dedicada a aquellos profesores de colegio, que marcaron para siempre mi vida futura y personal y, a mi mujer y mi hijo que son mi razón para vivir ese futuro. En particular a Benjamín León, mi profesor de Física de enseñanza media, John Mackenzie mi profesor de Teoría del Conocimiento y muy especialmente a Julio Blanco mi maestro en Inglés, Historia, Geografía y Matemáticas además de formarme en Ajedrez y Futbol. El profesor Blanco es, sin duda, el docente que más me ha marcado en mi vida, a través de su ejemplo de pasión y compromiso. Nunca olvidaré sus consejos, como aquel en que nos decía que un hombre debía formarse en todas las disciplinas, inspirándonos con Da Vinci, y no un experto en una sola cosa.

Agradezco a todos los alumnos de pedagogía, que participaron en las pruebas de este libro, aportando con sus comentarios y correcciones. Además, debo darle las gracias a un montón de personas que revisaron las versiones preliminares de este libro, en especial a Ernesto San Martin y Giovanna Tichione cuyos aportes me forzaron a reescribir gran parte del libro. A Mónica Celis por su exhaustiva revisión final del texto.

No puedo dejar de mencionar a Patricio Felmer y Salomé Martinez por su confianza y apoyo al encomendarme, esta enorme responsabilidad, en su hermoso proyecto. Finalmente agradezco al profesor Raul Gouet que me formó en probabilidades en la Universidad de Chile y que probablemente encontrará influencias de su estilo de enseñanza en este texto.

Es mi esperanza que este texto sea, un aporte real para nuestros profesores en esta área de las matemáticas que es tan especial para mí. Esta monografía ha sido escrita pensando en ellos o en cualquier lector cuyo interés sea, además de aprender por si mismo, enseñar probabilidades básicas. No hay que olvidar que el objetivo de enseñar probabilidades, no es sorprender ni entretener ni agudizar la mente, sino manejar un tópico desafiante de las matemáticas que tiene una inmensidad de aplicaciones en nuestro día a día. Escribir un texto orientado a docentes no es simplemente cambiar el contexto de los ejemplos de un texto clásico de probabilidades.

Si tuviera que resumir el espíritu de esta monografía en un sola frase, tendría que tomar prestadas las palabras de Pierre Simon Laplace que traduzco libremente del francés:

“En el fondo, la teoría de probabilidades no es sino sentido común reducido a cálculos”.

Es muy frecuente encontrar una gran cantidad de cursos de probabilidades que se podrían titular *Paradojas y Sorpresas de la Probabilidad* o bien *Juegos de Ingenio “para” probabilidades* que es como servir sólo el postre de una gran comida. El énfasis en los ejercicios utilizados para la enseñanza de las probabilidades debiera centrarse en construir intuición y no en paradojas engañosas. Es una gran tentación utilizar los problemas más rebuscados para ejemplificar los razonamientos erróneos, con la esperanza de, lograr de esta forma, iluminarlos. Sin embargo, el gran riesgo en persistir con esta práctica es que el lector se convenza de su completa incapacidad y en su desesperación desista de intentar comprender y se limite a aprender como resolver los problemas.

Por otro parte, es cierto que ejemplos contraintuitivos y paradojas tienen el efecto de captar la atención al plantear un desafío. Esto es un excelente camino para aprender y apreciar las habilidades de explorar, reflexionar y razonar. En general, los estudiantes tienden a aceptar lo que se les enseña aún en los casos en que esto entre en conflicto con sus intuiciones previas. Esto es gravísimo porque cuando los alumnos tengan que tomar un juicio de naturaleza probabilista fuera del salón de clase, lo *más probable* es que utilicen su intuición por sobre sus estudios formales.

En este trabajo no vamos de dejar de lado estas paradojas y ejemplos contraintuitivos, pero pretendemos utilizarlos sabiamente. Para citar a Laplace “Una de las grandes ventajas del cálculo de probabilidades es que nos enseña a desconfiar de nuestra primera impresión” y “La mente, como el sentido de visión, tiene sus ilusiones, tal como el tacto corrige estas últimas, a través de cálculos correctos se corrigen las primeras”.

Un estudiante debe construir su conocimiento de manera activa más que recibirla pasivamente de su entorno. Para lograr esto, los textos que utilizamos para formar a nuestros profesores debieran tener el mismo enfoque. Se debe entregar al docente más responsabilidad en su propia formación y ayudarlo a desarrollar pensamiento crítico y creativo. En una primera lectura pareciera que en este enfoque el rol del docente pasa a ser secundario lo que es un gran error. Su rol de hecho es más importante y complejo que antes, puesto que debe reconocer las creencias del alumno y ayudarlo a profundizar en ellas para que obtenga una mejor comprensión. Es entonces un *facilitador del conocimiento* y como tal, además de estar en contacto e interactuar con sus alumnos debe tener una solida formación en la disciplina que enseña.

El lector de este texto podrá apreciar que se dedican muchas de sus páginas a entregar diferentes representaciones y herramientas cuyo objetivo directo no es lograr desarrollar la intuición y ejemplificar conceptos, sino permitir al docente desarrollar su creatividad y ayudarlo a transmitir lo aprendido a sus alumnos. Los textos clásicos de probabilidades han sido escritos principalmente para ingenieros y economistas, en el último tiempo el área de la salud también ha generado una gran cantidad de obras pero en general más enfocadas a la Estadística. Si bien es cierto los contenidos pueden ser los mismos, la manera de presentarlos es por completo distinta. Nunca ha dejado de sorprenderme que, fuera del campo de los ingenieros, se considere valioso utilizar el concepto de Centro de Masa para aclarar el concepto de Valor Esperado de una Variable Aleatoria.

Enseñar probabilidades no puede limitarse solo a enseñar estructuras conceptuales y herramientas para la resolución de problemas. Es necesario desarrollar maneras de razonamiento y un sistema robusto de desarrollo de intuición. El razonamiento probabilístico es diferente del lógico o del razonamiento causal. Las cosas ya no son verdaderas o falsas siempre y la causalidad se puede invertir (Regla de Bayes).

Lo que normalmente debiera considerarse una gran ventaja, se ha convertido en un gran problema en la enseñanza de las probabilidades. Me refiero a su gran conexión con el mundo real. En nuestro propio lenguaje, expresiones como “independiente”, “más probable” y otras tienen una interpretación en castellano que no siempre corresponden con la definición formal que tienen en el contexto de probabilidades. Personalmente aprendí, con gran sufrimiento inicial y alivio final, que para un médico decir que una enfermedad es probable sólo significa que no puede descartarla como posible.

Existe un gran dilema en la enseñanza de las probabilidades respecto al nivel de formalismo que se requiere para enseñarlo. Un enfoque demasiado informal en probabilidades es una receta segura para el desastre. Es extremadamente fácil entregar diferentes argumentos que llevan a resultados diferentes en apariencia todos correctos. Al igual que en otras disciplinas, es un hecho que la representación que se utiliza para resolver un problema en probabilidades juega un rol fundamental llegando incluso a transformar un problema aparentemente complejo en algo obvio (no vamos a usar el término trivial). Por otro lado, el formalismo mínimo para solucionar por completo todos estos problemas es gigantesco y bastante más complejo de lo que se podría pensar. No en vano esto le tomó a la humanidad bastante tiempo como veremos más adelante.

El método actual de comenzar a enseñar probabilidades, curiosamente, repite los errores históricos del desarrollo de la Teoría de Probabilidades. Se genera una confusión total entre la Combinatoria y las Probabilidades al considerar el concepto intuitivo de base de la noción de probabilidad como el cociente entre casos favorables y casos posibles. Esto no es justo para ninguna de las dos disciplinas y en el caso de probabilidades es el principal causante de la gran mayoría de los ejemplos contraintuitivos y paradojas de la teoría. Aun si es cierto que, muchas de las propiedades básicas de la combinatoria permiten comprender de manera intuitiva algunas propiedades básicas de las probabilidades, muchas propiedades de la combinatoria que no son ciertas en probabilidades causan grandes problemas también. No pretendemos descartar este enfoque, pero no consideramos que sea un buen punto de partida para construir la intuición en probabilidades. Vamos a dar muchos ejemplos para justificar esta decisión y mostraremos las limitaciones de este paralelo entre combinatoria y probabilidad.

En este texto vamos a incorporar el formalismo de manera gradual y priorizaremos el desarrollo el desarrollo de la intuición. Para efectos de representación introduciremos elementos lúdicos, gráficos, concretos e intuitivos como nuestras herramientas principales. El primero de estos elementos serán discos con una cantidad finita de regiones distinguibles en los que el azar actúa como una aguja que gira sobre este disco. Este objeto tiene un atractivo natural para niños y jóvenes por lo que es ideal para introducir ejemplos entre ellos. El azar en la ocurrencia de un evento aleatorio estará representado por el giro de la aguja en un disco y el resultado de esta ocurrencia será la etiqueta que corresponde a la región del disco donde la aguja se detuvo. Con esta herramienta vamos a representar lo que formalmente se conoce como Variables Aleatorias Discretas y deduciremos de manera intuitiva las propiedades básicas de la Teoría de Probabilidades. Esto postergará la axiomática conjuntista de probabilidades hasta un momento más propicio en que se hayan desarrollado las habilidades intuitivas y la capacidad de modelar y simular. Permitirá además construir ejemplos y resolver problemas complejos con nuestras propias manos, con la intuición y sentido común como nuestra principal herramienta. Muchos de estos problemas serán retomados en el texto una vez introducida la axiomática, para ser resueltos nuevamente con este nuevo lenguaje. Una ventaja adicional consiste en realizar simulaciones en las que se integren elementos de geometría, que usualmente no se utilizan en probabilidades y que en general son mejor comprendidos en matemáticas.

Este texto también contiene los elementos básicos de un curso de combinatoria aunque limitado a aquellos elementos que resultan útiles para la teoría de probabilidades. Utilizando urnas y bolitas vamos a deducir las propiedades básicas de la combinatoria, entregando al mismo tiempo herramientas accesibles a cualquier docente que necesite enseñarlo a sus alumnos. Estas representaciones de disco, bolita y urna hacen natural, ameno e intuitivo el paralelo entre combinatoria y probabilidad y, al mismo tiempo. muestran las limitaciones de éste.

Capítulo 1: Introducción

1.1 Historia

Es importante conocer, al menos brevemente, la historia de las probabilidades, para poder aprender de sus lecciones antes de iniciarse en su estudio. Lo que se pretende no es una descripción exhaustiva, sino más bien resaltar los puntos que son importantes para la comprensión de la filosofía y la metodología que se va a desarrollar en este trabajo.

Es casi un concenso en nuestra cultura asociar las raíces de la teoría de probabilidades al nacimiento de los juegos de azar hace más de cuatro siglos. Las apuestas, de una manera u otra son tan antiguas como la propia humanidad. Según cuenta la historia, los chinos realizaban ya juegos de azar organizados por el Estado hace 2000 años. Durante la época del Imperio Romano, tuvieron gran auge los juegos de azar. Las apuestas deportivas eran parte de la vida de los romanos, incluso el propio Julio César apostaba en los acontecimientos que se celebraban en los circos romanos.

Figura 1.1. Romanos jugando a los dados

La lista de adictos a los juegos de azar en la historia es larga y extensa e incluye a emperadores romanos como Augusto y Claudio, literatos españoles como Góngora y Argote, y también rusos como Lermontov y Dostoievsky. Este último inclusive utiliza su trágica adicción en su obra clásica *El jugador*. Es claro entonces que la atracción hacia los juegos de azar es una constante en la historia que toca a todos los niveles sociales e intelectuales. De hecho en España en un estudio del 2007 se estima que el 2 % de su población es adicta a los juegos de azar (ludopatía) y alrededor de un 60 % de estos son menores de 30 años.

Esto justifica y, más aún, nos incentiva a mantener el enfoque lúdico como herramienta docente para la enseñanza de las probabilidades, especialmente para jóvenes.

Si nos remontamos a Aristóteles, este distingue tres tipos de causas o principios en la existencia, movimiento y posesión de uno u otro rasgo, propiedad o característica de los seres:

1. por azar: algo puede existir y ocurrir como consecuencia del azar: los llamados seres deformes o "monstruos de la naturaleza", la piedra que cae y que accidentalmente rompe una rama, etc...
2. por arte o técnica, como ocurre con cualquiera de nuestras máquinas y las cosas que ellas hacen
3. por naturaleza, como los cuatro elementos, las plantas, los animales y sus partes.

Durante la edad media hubo una gran actividad científica y artística en Oriente y el nombre de azar parece haber venido desde Siria a Europa. La flor de zahar, que aparecía en los dados de la época podría ser el origen.

Sin embargo, el enfoque de la Teoría de Probabilidades que utilizamos hoy es mucho más reciente. Alrededor de 1640 en la sociedad francesa, los juegos de azar comenzaron a ser extremadamente populares y no sujetos a restricciones legales. Cada vez fueron surgiendo juegos más complejos con cartas, dados y otros artefactos y además se apostaban mayores sumas de dinero por lo que las casas de juego consideraron una necesidad el desarrollo de algún sistema para predecir sus ganancias.

Es por esto que los ejemplos clásicos para explicar la Teoría de Probabilidades tienen que ver con los juegos de azar. Sin embargo, la utilidad actual de las probabilidades va mucho más allá que resolver juegos de salón y para ser el alma de las fiestas (al menos en los salones en la sociedad francesa del siglo XVII).

El primer libro sobre Teoría de la Probabilidades parece ser *De Ludo Aleae* de Girolamo Cardano (1501 - 1576) publicado recién en 1663 que está básicamente dedicado al juego de dados.

Se atribuye el origen de la formulación de la primera teoría de las probabilidades a una disputa entre apostadores en el año 1654, entre los cuales se encontraba Antoine Gombaud el Chevalier de Méré caballero y asiduo apostador. El Chevalier se contactó con Blaise Pascal, conocido matemático francés, para que le ayudara a discernir una aparente contradicción en un popular juego de dados. El juego consistía en lanzar un par de dados 24 veces. La pregunta era si valía la pena apostar a que salía un doble seis, al menos una vez, en los 24 lanzamientos. La experiencia indicaba que si era conveniente y por el contrario sus cálculos le decían que no. Aparentemente este y otros problemas planteados por el Chevalier, llevaron a Pascal a mantener un fluido intercambio de correspondencia con Pierre de Fermat otro matemático francés.

Este intercambio se considera como el origen de los principios de la Teoría de las Probabilidades. Sin embargo, debe tenerse en cuenta que, Pascal y Fermat no desarrollaron mucho el campo. Sería Christian Huygens en 1657, quien, tras conocer el tema a través de la correspondencia entre Pascal y Fermat, escribiría *De Ratiociniis in Ludo Aleae*, el primer libro sobre el particular. Otras figuras que continuaron el desarrollo de ese campo matemático serían también Abraham de Moivre y Pierre-Simon Laplace.

Durante el siglo 18 el interés por aplicar matemática a los juegos de azar se expandió rápidamente, siendo Jakob Bernoulli y Abraham de Moivre dos de los principales precursores de este movimiento. Sin embargo, en esta época las probabilidades ya eran aplicadas fuera del contexto de los juegos de azar. El Marqués de Condorcet (1743-1794) líder político durante la Revolución Francesa estaba interesado en aplicar Teoría de probabilidades a economía y política. El realizaba cálculos para obtener la probabilidad de que un jurado, que decidiera por mayoría, tomara la decisión correcta si cada miembro del jurado tenía la misma probabilidad de tomarla independientemente.

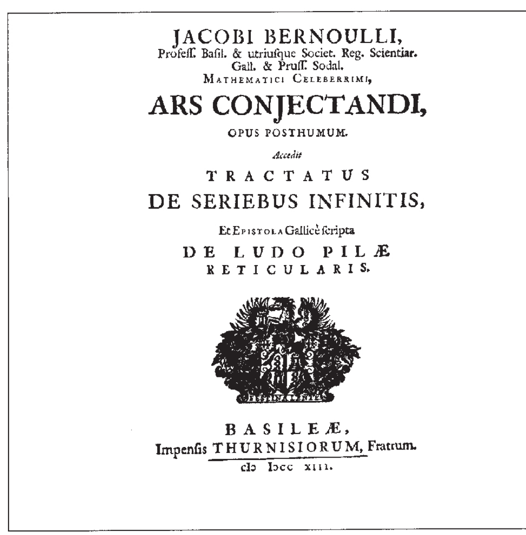

JACOBI BERNOULLI,
Profeſſ. Baſil. & utriuſque Societ. Reg. Scientiar.
Gall. & Pruſſ. Sodal.
MATHEMATICI CELEBERRIMI,
ARS CONJECTANDI,
OPUS POSTHUMUM.
Accedit
TRACTATUS
DE SERIEBUS INFINITIS,
Et EPISTOLA Gallicè ſcripta
DE LUDO PILÆ
RETICULARIS.
BASILEÆ,
Impenſis THURNISIORUM, Fratrum.
clɔ Iɔcc xiii.

Figura 1.2. Portada de la obra de J. Bernoulli

En 1812 Pierre de Laplace publica su obra titulada *Théorie Analytique des Probabilités* en la que introduce nuevas ideas y técnicas al cálculo de las probabilidades. Antes de Laplace la Teoría de Probabilidades era tan solo un análisis matemático de los juegos de azar. Así, Laplace extendió definitivamente el uso de las probabilidades fuera del ámbito de los juegos de azar. La Teoría de Errores y la Mecánica Estadística son algunos ejemplos.

En 1894, Karl Pearson analizó un gran número de resultados de una determinada ruleta no justa (con distribución no uniforme) y sugirió utilizar los casinos como un laboratorio aplicado de Teoría de Probabilidades para realizar experimentos, lo que lo condujo a descubrir la prueba Chi-Cuadrado de independencia.

Una de las mayores dificultades para desarrollar una teoría de las probabilidades, es obtener una definición que sea a la vez matemáticamente consistente y con la simpleza necesaria para ser aplicada a una gran variedad de situaciones. Esta búsqueda tomó más de tres siglos y estuvo marcada por mucha controversia.

De hecho, grandes matemáticos cometieron errores, que en nuestros días se considerarían de principiante. Por ejemplo, Jean le Rond D' Alembert (1717-1783) (uno de los autores de la famosa Enciclopedia Francesa) al responder la pregunta; ¿con que probabilidad una moneda que se tira dos veces, por lo menos una vez cae en sello? entregó como respuesta $\frac{2}{3}$. El matemático consideraba que hay únicamente tres resultados posibles, dos sellos, 1 sello y una cara o 2 caras y un caso no favorable. D' Alambert no tuvo en cuenta que los tres resultados posibles no eran igualmente probables!! (La respuesta correcta es $\frac{3}{4}$ como aprenderán fácilmente los lectores de esta monografía). Hemos resaltado este ejemplo, porque este tipo de errores provienen de plantear la enseñanza de las probabilidades como una aplicación de la combinatoria, al considerar como la fórmula de probabilidad el cociente de favorables sobre posibles, que es algo que pretendemos evitar en este texto.

Esta génesis es la que habitualmente se considera como el origen de la teoría de probabilidades, pero lo cierto es que en esa misma época, en la Madre Rusia florecía una escuela matemática que finalmente haría muchas de las contribuciones fundamentales a la Teoría de Probabilidades, que está plasmada en muchos de los nombres que utilizamos actualmente para grandes resultados. El primer libro de probabilidades en ruso fue escrito por Viktor Buniakovski (1804-1889) y no es posible dejar de mencionar a Pafnuty Tchébychev (1821-1894) que, además de su famosa desigualdad, generalizó la Ley de Grandes Números y a Sergei Michailovich Liapunov (1857-1918) creador de la primera demostración del Teorema Central del Límite para Variables Aleatorias Independientes. Andrei Andreyevich Markov (1856-1922) extendió los dominios de aplicación de la Ley de los Grandes Números y el Teorema Central del Límite y fue el creador de la noción de Proceso Estocástico.

Para finales del siglo XIX la teoría de probabilidades y el concepto de azar se habían transformado en temas polémicos por su alto potencial matemático. Muchos matemáticos volvían a plantearse preguntas fundamentales. Poincaré en su libro *Science et Méthode* planteaba que la noción de azar no se debe tanto a nuestra ignorancia como a una falta de soporte empírico. Estas reflexiones estaban motivadas por la discusión fundamental entre determinismo e indeterminismo levantada por la crisis de la Física de los primeros años del siglo XX.

En el Segundo Congreso Internacional de Matemáticas (París 1900) David Hilbert planteó 23 problemas a sus colegas matemáticos. Entre esos problemas, el sexto, solicitaba encontrar una base axiomática que permitiese deducir todas las teorías físicas y los fenómenos aleatorios o dependientes del azar. Textualmente (aparte de su traducción): "Las investigaciones sobre los fundamentos de la geometría sugieren el problema: Tratar de la misma manera, por medio de axiomas, aquellas ciencias físicas en las cuales las matemáticas juegan una parte importante; en primer lugar están la teoría de probabilidades y la mecánica."

Muchos matemáticos destacados como Kolmogorov y von Neumann dedicaron gran parte de su carrera académica al estudio de las probabilidades. Finalmente en 1933, fue la monografía del matemático ruso Andrei Kolgomorov (traducida al inglés en 1950), titulada *Foundations of Probability Theory*, la que nos dio la axiomática que es la base de la teoría moderna de las probabilidades.

Todo este desarrollo de modelos matemáticos del azar ha consolidado la Teoría de Probabilidades y sus disciplinas relacionadas (teoría ergódica, teoría de la información, teoría de procesos estocásticos entre otras) que permiten la construcción del Análisis Estocástico en el curso del presente siglo.

Es sólo después de la mitad del siglo XX, y a partir de los años 1970, que la Teoría de Probabilidades se enseña como parte de todas las carreras profesionales vigentes hoy en día. En la actualidad su conocimiento se exige prácticamente en todos los estudios de postgrado de Ingeniería, Ciencias de la Salud e incluso en Ciencias Sociales.

En gran parte del mundo las probabilidades forman, incluso, parte del programa de estudios tanto en Educación Básica como Media. Sin embargo, son pocas las carreras de Pedagogía en Matemáticas en nuestro país que tienen una formación sólida en probabilidades. En muchos casos existe un sólo curso, y a veces electivo, en el área de probabilidades que en la mayoría de las veces también se usa para enseñar estadística.

Capítulo 2: ¿Qué es el Azar?

Existen libros completos que se dedican a examinar la buena definición de azar y sus consecuencias epistemológicas y filosóficas. En nuestro caso vamos a definir el azar de la manera más simple y útil que nos sea posible. De manera general vamos a cuantificar el azar según la *falta de información*. Es decir, mientras menos información se tenga mayor será el azar. En palabras de Henri Poincaré: "El azar es una medida de nuestra ignorancia."

Un *Evento Aleatorio* ($\mathcal{E}$) es cualquier operación, ya sea real o ficticia, que entrega uno o más resultados distinguibles y conocidos. A pesar del nombre, un Evento Aleatorio no tiene porque tener una naturaleza aleatoria en el sentido castellano del término. La real academia define aleatorio como "Perteneciente o relativo al juego de azar" o bien "Dependiente de algún suceso fortuito" y por otro lado define azar como "Casualidad, caso fortuito".

La falta de información no necesariamente proviene exclusivamente del evento. El observador es la otra causa principal y, por tanto, el evento en sí podría ser completamente no aleatorio, siendo el azar exclusiva responsabilidad del observador. Una primera consecuencia importante de esta observación, es que nuestra representación de un Evento Aleatorio puede evolucionar al incorporar nueva información del Evento Aleatorio o la percepción y/o cambio del observador.

En vista y considerando lo anterior, los Eventos Aleatorios no aparecen sólo en el ámbito de los juegos de azar y de hecho su utilidad ha demostrado ser fundamental en diversas áreas del conocimiento. En palabras del gran Laplace: "Es un hecho destacable que una ciencia que empezó analizando juegos de azar acabe convirtiéndose en el más importante objeto del conocimiento humano."

Virtualmente, cualquier situación que se pueda representar como una lista de resultados posibles es un Evento Aleatorio. Invito a los lectores a pensar en una situación de interés que no pueda ser representada de este modo. Como decía el filósofo griego Demócrito de Abdera: "Todo lo que existe en el universo es fruto del azar y de la necesidad."

Más que una disciplina en particular, es una nueva manera de analizar cualquier fenómeno y por tanto no tiene limitaciones en su campo de aplicación. Por otro lado, es verdad que su utilidad es limitada y está subordinada a la habilidad de cuantificar y utilizar la falta de información disponible. Un modelo que no tiene utilidad no tiene sentido. De otro modo no podríamos defendernos de los dichos de Bertrand Russell: "¿Cómo osamos hablar de leyes del azar?, ¿no es, acaso, el azar la antítesis de cualquier ley?."

Nuestra visión contrasta un poco con las ideas establecidas sobre cuándo se debe utilizar el cálculo de probabilidades. Como Descartes decía: “Es una verdad muy cierta que, cuando no esté a nuestro alcance determinar lo que es verdad, deberemos seguir lo que es más probable.”

Es decir, apliquemos las probabilidades cuando no tengamos otra alternativa. Es un hecho que en nuestros días esta posición es insostenible. Con la aparición de la mecánica cuántica, por ejemplo, hubo gran oposición a la idea de incorporar cálculos probabilistas. Como dijo Einstein en una de sus citas más racistas, a mi parecer: “Dios no juega a los dados.”

La respuesta de Stephen Hawking a esta cita me parece lapidaria y aplastante: “Dios no solo juega a los dados sino que a veces los lanza donde no los podemos ver.”

El problema entonces no es modelar como un Evento Aleatorio sino cómo obtener la información cuantitativa necesaria para que este enfoque nos sea de utilidad.

Quisiera terminar citando a un gran científico nacional, Servet Martínez: “El completo azar a nivel microscópico, es lo que nos permite tener leyes deterministas a nivel macroscópico.”

2.1 Presentación

En este capítulo vamos a formalizar el concepto de *Azar* como *falta de información*. Cualquier experimento que permite obtener una lista de resultados posibles, sin saber cual de ellos ocurrirá en cada realización, se considera como un *Evento Aleatorio*. Entonces, cualquier experimento se puede considerar aleatorio pero, la falta de información es relativa al observador y podría variar de una realización a otra para él.

Si consideramos el caso de listas finitas de resultados posibles (que llamaremos *Caso Discreto*) lo que primero distingue a dos representaciones de Eventos Aleatorios es el conjunto de resultados posibles. Si los resultados posibles son los mismos, entonces el segundo y último factor de distinción es la estimación que se tiene de cuán posible es obtener cada uno de ellos. Dos Eventos Aleatorios representados de esta manera, no son distinguibles, si tienen los mismos resultados posibles y es tan factible obtener cada uno de ellos.

La representación discreta de un Evento Aleatorio se puede efectuar en un Disco o Ruleta con una cantidad finita de regiones con etiquetas que representan dichos resultados. La proporción de arco de cada región es una medida de cuán posible es obtener ese valor en particular. Una realización del Evento Aleatorio será una aguja que gira y el resultado obtenido será el indicado por la región donde se detuvo la aguja.

Este enfoque permite construir un simulador del fenómeno de manera concreta y simple. Sin embargo, para realizar ciertos cálculos, veremos que es mejor otra representación como una tabla de dos filas en que los resultados posibles están en la primera fila y la proporción de perímetro de la región asociada está en la segunda fila. Al considerar la combinación de más de un Evento Aleatorio en el caso discreto, veremos que resulta de utilidad representarlos como un árbol. Estudiaremos como construir cada una de estas representaciones a partir de cualquier otra o del evento original. Con estas herramientas resolveremos diversos problemas y deduciremos la mayoría de las propiedades básicas de la Teoría de Probabilidades para el caso discreto.

2.2 Espacio Muestral

Frente a un Evento Aleatorio, lo primero a definir es la lista de resultados posibles distinguibles por el observador. A esta lista la llamaremos *Espacio Muestral*. Mientras mayor sea la cantidad de elementos del Espacio Muestral mayor será la falta de información (es decir el azar). En el caso de una cantidad finita de resultados posibles se habla de un *Espacio Muestral Discreto*.

Los ejemplos clásicos de Espacios Muestrales Discretos provienen de los juegos de azar, por ejemplo, cara o sello como valor de la cara superior al lanzar una moneda. El tiempo de vida de una ampolleta (o una persona, en una versión más trágica) resulta no ser discreto puesto que puede ser cualquier número real positivo.

En una definición más formal, el Espacio Muestral se representará como un conjunto y se utilizará toda la estructura de la Teoría de Conjuntos para dar definiciones, demostrar propiedades y realizar cálculos.

Un mismo Evento Aleatorio puede dar pie a distintos Espacios Muestrales que, inclusive, podrían ser en algunos casos discretos y en otros no. Esto resulta ser extremadamente importante y muchas veces no queda tan claro. Por ejemplo, un clásico de los juegos de azar es lanzar un dado. El Evento Aleatorio es el lanzamiento del dado. Se asume que es evidente que lo relevante de este Evento es el valor que se observa en la cara superior del dado al dejar de rodar. Es decir, el Espacio Muestral Discreto que está formado por los números del 1 al 6. Esto porque en nuestra cultura es lo que habitualmente interesa en los juegos de dados. Es posible que en otras culturas y/o aplicaciones se considere otra característica del experimento, que podría ser inclusive un Espacio Muestral infinito.

Determinar el Espacio Muestral correcto puede resultar más difícil de lo que parece. Del ejemplo anterior queda claro que el Evento por sí solo no permite definir el Espacio Muestral. Es necesario determinar el interés en el Evento Aleatorio para poder definir el Espacio Muestral adecuado. Por ejemplo, en el caso del dado, ganar el juego depende del número de la cara superior, esos números formarán el Espacio.

Es fundamental entender la importancia de seleccionar un Espacio Muestral. Este es el primer paso para *Modelar* nuestro Evento Aleatorio y para esto se debe tener un objetivo claro. No se modela por el puro placer intelectual de hacerlo. Una manera tradicional de fijar el objetivo consiste en decidir qué tipo de preguntas son las que se quiere responder. Volviendo a nuestro ejemplo del lanzamiento del dado. Es muy diferente modelar pensando en preguntas del estilo, ¿cuál es la probabilidad de obtener un seis? que modelar pensando en, ¿cuál es la probabilidad que el dado ruede por más de 5 segundos?.

Por desgracia, tener claro el objetivo tampoco define por completo el Espacio Muestral. Un ejemplo clásico, que mencionamos en la historia de las probabilidades, proviene de considerar el Evento Aleatorio de lanzar una moneda dos veces. Podemos considerar como resultados posibles el número de caras obtenidas lo que nos da 0, 1 o 2 o bien la lista de resultados que nos da cara-cara, cara-sello, sello-cara y sello-sello. Ambos Espacios Muestrales son adecuados para responder preguntas, por ejemplo, respecto a cuales son las posibilidades de obtener dos caras seguidas. Veremos, sin embargo, que al momento de cuantificar cuán probables son los resultados posibles, el que aparentemente es más directo para resolver el problema, es decir, el número de caras obtenidas, no es el más adecuado puesto que es más natural cuantificar los resultados posibles para el otro caso.

2.2.1 El Problema de las 12 Monedas

Es sorprendente la cantidad de información que se puede obtener tan solo considerando la noción de Espacio Muestral; Por ejemplo, consideremos el Evento Aleatorio con una bolsa con 12 monedas de las cuales una es falsa (la única manera de distinguirla es que su peso es significativamente distinto del de las otras). Si se nos da una balanza de contrapeso, la pregunta a resolver es la cantidad mínima de veces que podemos utilizar la balanza para determinar la moneda falsa. Este es un problema cuyo origen se pierde en la historia, pero que resulta ser un clásico como problema de ingenio; Ha sido resuelto de múltiples maneras por ilustres personajes entre los que destacamos a Martin Gardner.

El Espacio Muestral natural para encontrar la moneda falsa son 12 resultados posibles, uno por cada moneda. Si utilizamos la balanza una sola vez, sólo obtenemos 3 resultados distinguibles que corresponden a cuando el lado izquierdo de la balanza pesa más, menos o igual que el lado derecho. Es claro que el número de posibilidades no alcanza para obtener una respuesta a nuestro problema, no es posible decidir entre 12 posibilidades utilizando 3. Si usamos la balanza 2 veces, el número de resultados posibles será $3 \cdot 3 = 9$, que todavía no es suficiente. Al usarla 3 veces, obtenemos $3 \cdot 3 \cdot 3 = 27$. Por lo tanto, la conclusión es que al menos se debe utilizar la balanza 3 veces para encontrar la moneda falsa.

Si sólo utilizamos 2 resultados posibles por cada uso de la balanza, por ejemplo, si ambos lados pesan igual o distinto necesitaremos utilizarla 4 veces en vez de 3!!.

Es importante aclarar que no hemos resuelto el problema (ni estamos cerca de hacerlo) pero sí, hemos sido capaces de determinar que la balanza debe ser utilizada al menos 3 veces en un caso y 4 veces en el otro, considerando sólo el concepto de Espacio Muestral como herramienta. Si no fuese posible resolver el problema con 3 usos de la balanza esto, quiere decir que la información útil por cada uso de la balanza no es 3 sino tan solo 2.

2.3 Probabilidad sobre el Espacio Muestral

Recordemos una vez más que un Evento Aleatorio no tiene porque tener una naturaleza lúdica o aleatoria. El azar muchas veces se manifiesta exclusivamente por la falta de información del observador. La primera diferenciación entre dos o más Eventos Aleatorios corresponde al Espacio Muestral asignado. Sin embargo, aún si dos Eventos Aleatorios tienen el mismo Espacio Muestral esto no los identifica como uno solo.

El siguiente paso para determinar un Evento Aleatorio corresponde a cuantificar cuan factibles son cada uno de los resultados posibles. En el caso de Espacios Muestrales Discretos, la tradición corresponde a asociar a cada resultado un valor numérico positivo entre 0 y 1. Mientras más pequeño el valor asociado, menos probable es que ese resultado ocurra en una realización del Evento Aleatorio. Entonces el valor 0 indica que este resultado es casi imposible (aunque no imposible) y el valor 1 lo contrario, es decir, que es casi imposible que no ocurra.

Por ejemplo, el resultado de un partido de fútbol o de una elección no es aleatorio después de finalizado el partido o de realizada la votación. Sin embargo, para un observador que no pudo ver el partido o antes de que los votos sean contados, esto si será un Evento Aleatorio. El azar en estos casos depende exclusivamente del observador y por tanto, puede cambiar de acuerdo a la información de la que dispone dicho observador.

Un observador completamente desinformado y sin ideas preconcebidas, naturalmente asignará la misma opción de ganar a cada candidato o equipo. Un hincha de corazón al igual que un miembro fanático de alguna colectividad política, asignará una posibilidad nula a que su favorito pierda. Un observador más informado y realista podrá considerar las estadísticas recientes disponibles. Esto podría incluir, en el caso del partido de fútbol, los resultados de los últimos partidos jugados de cada equipo o de los últimos encuentros entre estos dos equipos en particular en toda su historia. En el caso de una elección se podría considerar las encuestas previas a las elecciones o la cantidad de inscritos en los partidos que apoyan a cada uno de los candidatos. Las fuentes de información para cuantificar los resultados posibles son infinitas y en algunos casos pueden parecer completamente absurdas en nuestra cultura (por ejemplo visiones, sueños o predicciones astrales). Validar la confiabilidad, o veracidad, de la cuantificación escogida no corresponde a este trabajo, dado que eso corresponde más al área de la Estadística. Lo único que nos importa de la fuente de información es que se pueda traducir en una cuantificación de los resultados posibles del Espacio Muestral escogido.

Una vez decidida la asignación inicial, corresponde observar para incorporar nueva información que permita mejorar el modelo de la mejor forma posible. Por ejemplo, el marcador del medio tiempo o una cuenta parcial de votos pueden modificar la manera de cuantificar los resultados posibles. Lo ideal es encontrar una manera de incorporar esta nueva información a la ya existente y no comenzar todo de nuevo. Es esta necesidad la que produce las reglas del cálculo de probabilidades que vamos a discutir en este trabajo. Como siempre digo, la idea de las matemáticas es lograr trabajar lo menos posible para lograr el objetivo. En general, pensar primero y calcular después, tiende a ser más productivo en este sentido que calcular adivinando con la esperanza de tropezar con la respuesta correcta. No hay que olvidar nunca que pretendemos guiarnos por el sentido común para traducirlo en cálculos.

En ciertas tribus africanas se tenía la creencia que tocar los tambores al amanecer era lo que causaba que el sol saliera cada mañana. El riesgo de que esto fuese posible aunque poco probable, impedía que dejaran de realizar este ritual día tras día para verificar esta hipótesis. En nuestra cultura esto parece ridículo, pero no lo es más que jugar a la ruleta rusa o comprar un boleto de lotería con la esperanza de ser millonario algún día.

En la literatura clásica, se reconocen tres estilos básicos de asignación inicial de probabilidad, que enunciamos a continuación. Lo cierto es que, en nuestro caso, nos vamos a concentrar en uno de ellos.

2.3.1 Enfoque a Priori

Comenzamos con el *Enfoque Clásico o A Priori* (del latín "lo que viene antes de"). Se define como "una cuantificación obtenida a través de un análisis lógico de la información disponible del Evento Aleatorio". Este trabajo pretende principalmente desarrollar las habilidades necesarias para poder utilizar todo el potencial de esta filosofía de cuantificar las probabilidades.

En el caso de absoluta falta de información sobre el Evento Aleatorio (aparte del Espacio Muestral) este enfoque nos lleva a asignar a cada resultado posible, la misma opción que cualquier otro (esto se justifica normalmente por simetría). En el caso discreto esto quiere decir que si tenemos n resultados posibles cada uno tiene una probabilidad de $\frac{1}{n}$. Es decir, la probabilidad de que alguna situación ocurra se reduce a la fórmula de *favorables sobre posibles*, que muchos autores llaman la *Regla de Laplace*. Esto, a veces se confunde, con la regla general de este enfoque y, en algunos libros, inclusive lo definen de esta manera. Como los lectores comprenderán a estas alturas, asumir esto como herramienta para desarrollar la intuición tiene consecuencias nefastas dado que se basa en la absoluta falta de información. Regresemos a un ejemplo, que entregamos en nuestro breve recuento histórico, para aclarar lo expresado. ¿Cuál es la posibilidad de que una moneda que se lanza dos veces, por lo menos una vez caiga sello?. Los resultados posibles pueden ser 3 o 4 según como razonemos. Si pensamos en el número de sellos obtenidos tenemos 0, 1 o 2 y por tanto, sin mayor información, respondemos una probabilidad de $\frac{2}{3}$. Pero si consideramos los 4 resultados "reales" tendremos 4, es decir cc, cs, sc y ss (donde c es cara y s es sello) por lo que, sin mayor información, respondemos $\frac{3}{4}$. Esto nos presenta una aparente paradoja, puesto que dos razonamientos correctos producen resultados distintos.

Lo cierto es que existe un razonamiento injustificado en la asignación de las probabilidades en el primer caso (no un error en seleccionar un Espacio Muestral de 3 resultados posibles!!). No es razonable asumir que 0, 1 o 2 sellos son igualmente probables, puesto que sólo un resultado permite no obtener sellos (2 caras seguidas) y; 1 sello se puede lograr con una cara y un sello (y al revés) y, 2 sellos (una vez más de una sola manera). Entonces las probabilidades de 0 y 2 sellos debieran ser la mitad de la probabilidad de obtener 1 sello. Es decir $\frac{1}{4}$ para 0 y 2 y $\frac{1}{2}$ para 1 sello si queremos suma total 1. Entonces la probabilidad de obtener al menos 1 sello coincide con la respuesta obtenida con el segundo Espacio Muestral de 4 resultados.

A juicio del autor, la gran mayoría de las paradojas, y las que más confusión producen entre los aprendices de probabilidades, son consecuencia de esta simplificación. Este tipo de paradojas son en realidad solo un resultado contraintuitivo si basamos nuestra intuición en la Regla de Laplace. Es lamentable constatar que este método se usa poco correctamente, a pesar de que tiene enormes ventajas en la comprensión y enseñanza de la disciplina y no es mucho más complejo de utilizar. Otra ventaja adicional que se obtiene al potenciar este enfoque es el desarrollo de habilidades de razonamiento lógico que son necesarias e inevitables para la cuantificación de los resultados posibles. A veces me pregunto si evitar el aprendizaje de estas habilidades con el loable objetivo de hacer el aprendizaje más fácil, fue lo que llevó a considerar la enseñanza solo con el caso equiprobable.

2.3.2 La Paradoja de Simpson

Un ejemplo famoso de estas falsas paradojas se conoce como la *Paradoja de Simpson*. El planteamiento lógico del problema es bastante simple, pero he preferido dar un versión real de aplicación de esta mala comprensión de las probabilidades.

En los años 70 la División de Admisión de la Universidad de Berkeley en el estado de california, realizó un estudio para comprobar un reclamo de prácticas discriminatorias en la admisión respecto al género (para detalles ver la revista *Science*, vol 187, Pag. 398-404, 1975, Bickel, Hammel y O'Connell).

La información disponible indicaba que 8442 hombres y 4321 mujeres postularon en admisión; El 44 % de los hombres fueron admitidos y solo el 35 % de las mujeres. Esto parecía apoyar el reclamo e indicaba una discriminación hacía el género femenino.

Sin embargo, el descomponer estos datos con un poco más de información, los investigadores (que eran dos estadísticos y un decano) se encontraron con resultados inesperados. Reemplazando los nombres de los departamentos por las letras A,B, C,D, E y F y los datos de admisión (total de alumnos y porcentaje admitido para cada género) para los 6 departamentos con más ingreso, se obtenía:

	Hombres		*Mujeres*	
Dep.	*Total*	*%Ad.*	*Total*	*%Ad.*
A	825	62	108	**82**
B	560	63	25	**68**
C	325	**37**	5938	34
D	417	33	375	**35**
E	191	**28**	393	24
F	373	6	341	**7**
Tot.	2691	45	7180	33

En este grupo más reducido los resultados apoyan de manera más aplastante el reclamo, puesto que el porcentaje de admisión de hombres es de un 45 % y de mujeres de un 33 %. Sin embargo, la conclusión al analizar cada departamento por separado es diferente. En efecto, existe una fuerte discriminación hacia el género masculino en el departamento A y hay un mayor porcentaje de admisión de mujeres en los departamentos B, D y F, siendo la diferencia en los departamentos C y E poco significativa a favor del género masculino. Encontramos entonces, dos razonamientos correctos, que llevan a conclusiones opuestas. Por lo tanto, tenemos una paradoja.

Para explicar de manera más sencilla esta situación consideremos un segundo ejemplo con datos reales. El mejor bateador de una temporada de béisbol, es aquel que tiene el mejor porcentaje de hits (golpes conectados) sobre el total de lanzamientos que ha jugado. Los rendimientos de los dos mejores bateadores en los años 1989 y 1990, Dave Justice y Andy Van Slyke fueron los siguientes:

	Dave Justice			Andy Van Slyke		
Año	Hits	Total	Hits/Total	Hits	Total	Hits/Total
1989	12	51	0,235	113	476	0,237
1990	124	439	0,282	140	493	0,284
Total	136	490	0,278	253	969	0,261

La paradoja en este caso consiste en constatar que el ganador de las temporadas 1989 y 1990, no es el ganador si se consideran las dos temporadas como una sola gran temporada, en cuyo caso el ganador resultar ser el otro bateador. Parece absurdo constatar que si en ambas temporadas el promedio de hits de un jugador es superior al de otro, en el promedio de ambas temporadas esto se invierte.

Esto nos parece contradictorio, solo por el hecho que en el caso equiprobable esto es imposible. El promedio de cada bateador se obtiene como el *promedio ponderado* de los promedios de ambas temporadas es decir:

Si los ponderadores fueran todos iguales, esto sería aritméticamente imposible puesto que:

$$
\boxed{\begin{array}{lllll}
0{,}278 = & 0{,}235 & \cdot\dfrac{51}{490} & +0{,}282 & \cdot\dfrac{439}{490} \\
\veebar & \barwedge & & \barwedge & \\
0{,}261 = & 0{,}237 & \cdot\dfrac{476}{490} & +0{,}284 & \cdot\dfrac{493}{969}
\end{array}}
$$

FIGURA 2.1. Paradoja de Promedios Ponderados

$$
\boxed{0{,}235 \leq 0{,}237 \text{ y } 0{,}282 \leq 0{,}284 \Rightarrow \frac{1}{2}0{,}235 + \frac{1}{2}0{,}282 \leq \frac{1}{2}0{,}237 + \frac{1}{2}0{,}284.}
$$

2.3.3 Enfoque Empírico

Proseguimos con el *Enfoque Clásico Empírico*, que consiste en cuantificar los resultados posibles, en el mismo espíritu de la Regla de Laplace. En esta situación se utiliza la información disponible, de realizaciones anteriores, del Evento Aleatorio. Se considera la opción de cada resultado posible, como el cociente entre las veces que el resultado ha ocurrido sobre el total de realizaciones del Evento. Este cociente usualmente se llama la *frecuencia relativa* de dicho resultado. En nuestro enfoque, este es el siguiente que incorpora información aparte del Espacio Muestral, considerando como nueva fuente de información la experiencia de diversas realizaciones. Esta manera de asignar probabilidades se explota más en el área de la Estadística que en nuestro contexto de aprendizaje de la Teoría Básica de Probabilidades.

Existe una justificación matemática formal, que explica y justifica, por qué esta asignación resulta ser adecuada para cierta clase de Eventos Aleatorios. Por desgracia, no tenemos en este instante el nivel de formalismo y conocimientos que se requieren. A pesar de ello algo podemos decir de manera informal y, apelando a la intuición del lector.

Primero, es evidente que esta información es más precisa y útil mientras mayor sea el número de realizaciones del Evento. En particular con 1 realización la frecuencia relativa solo puede ser 1 o 0 y con 2 solo $\frac{0}{2}, \frac{1}{2}$ o $1 = \frac{2}{2}$ y así sucesivamente. Es por tanto de utilidad disponer de una cantidad importante de realizaciones del Evento, respecto a la cantidad de resultados posibles. Recordemos el espíritu del análisis del problema de la moneda falsa. No será de mucha utilidad, tener 3 realizaciones de un evento, como fuente de información, si los resultados posibles son 12.

Segundo, este proceso tiene una tendencia a centrarse en torno a un valor (*converger*, en términos más matemáticos). En efecto, si consideramos n realizaciones con, digamos, k resultados favorables (es decir una frecuencia relativa de $\frac{k}{n}$) y, obtenemos información de otra realización la frecuencia relativa actualizada solo puede tomar los valores $\dfrac{k+1}{n+1}$ o $\dfrac{k}{n+1}$. Es decir, la frecuencia relativa siguiente difiere a lo más en $\dfrac{1}{n+1}$ (ver ecuación 2.1 para la deducción).

$$\left|\frac{k}{n}-\frac{k+1}{n+1}\right| = \frac{k\cdot(n+1)-(k+1)\cdot n}{n(n+1)} = \frac{n-k}{n(n+1)} \leq \frac{1}{n+1} \tag{2.1}$$

$$\left|\frac{k}{n}-\frac{k}{n+1}\right| = k\cdot\frac{(n+1)-n}{n(n+1)} = \frac{k}{n(n+1)} \leq \frac{1}{n+1}. \tag{2.2}$$

Dado que $\dfrac{1}{n+1}$ se hace más pequeño, cuando n aumenta, este cociente cambia cada vez menos y en muchos casos tiende a estabilizarse en torno a un *valor límite*. Hay una sutileza matemática que hace que nuestro argumento no garantice la convergencia, aunque intuitivamente nos parezca que debe ser así. No olvidemos que, la convergencia de esta secuencia de frecuencias (bajo hipótesis adicionales) es uno de los grandes Teoremas en probabilidades, llamado la *Ley de los Grandes Números*.

Por último, es claro que el orden en que incorporamos las realizaciones no es importante (es decir, son *intercambiables*), puesto que la frecuencia relativa se mantiene igual ante estos cambios. Esto quiere decir que, esto es sólo aplicable, cuando el orden de las realizaciones no cambia la información de cuán probable es que ocurran los resultados posibles en cada etapa. Esto se relaciona con el concepto de *independencia* que trataremos más adelante.

En general este tipo de enfoque se aplica en situaciones como múltiples lanzamientos de una moneda, dado u otro objeto lúdico; En esta situación parece aceptable asumir lo anterior. Sin embargo, veremos que una ligera modificación de las condiciones (ejemplo 3.1) nos lleva a conclusiones erradas. Entonces para determinar la probabilidad de que un jugador gane un partido contra una máquina por ejemplo este enfoque no es aplicable, a menos que supongamos que la experiencia adquirida por la persona en cada juego es nula.

Un razonamiento clásico y de gran polémica, utilizando información de realizaciones previas, lo realizó Pascal al enunciar su famosa *Regla de la Sucesión* que se puede enunciar, informalmente, como sigue; si lo único que se sabe de un evento es que ha ocurrido n veces seguidas, la probabilidad de que ocurra de nuevo será $\dfrac{n+1}{n+2}$ (esta deducción no es tan directa y no la vamos a explicar). Con esto Pascal estimó que la probabilidad de que saliera el Sol mañana era $\dfrac{1826213+1}{1826213+2} = 0{,}9999995$. En nuestros días esta estimación está aún más cercana a 1 todavía!.

2.3.4 Enfoque Subjetivo

El último caso a considerar se denomina *Probabilidad Subjetiva* que se define como "la cuantificación que se obtiene por el juicio subjetivo de un observador determinado"; en este caso no es necesario ningún tipo de lógica, o justificación, en la asignación. Al igual que el Enfoque Empírico éste resulta de mucha utilidad en el área de la Estadística, porque permite incorporar información de "expertos" que no tienen necesariamente un razonamiento lógico o empírico que justifique su decisión.

2.4 Caracterizando Eventos Aleatorios

En nuestro contexto un Modelo de un Evento Aleatorio consiste en tan solo dos elementos.

1. Espacio Muestral.
2. Cuantificación de cuán posible es cada uno de los resultados respecto a los demás.

En el caso de Espacios Muestrales Discretos la cuantificación resulta ser bastante simple. Consiste simplemente en cuantificar cuán factible es cada uno, de los valores posibles, respecto a los demás.

2.4.1 Espacios Muestrales Infinitos

A pesar de que esto no será tratado en este texto, es importante aclarar que cuantificar la probabilidad de cada uno de los resultados posibles, en el caso de un Espacio Muestral infinito no permite construir un modelo consistente. Por ejemplo, consideremos un número natural cualquiera, al azar. Si no tenemos más información, todos los números debieran tener la misma probabilidad. Ahora bien, dado que la suma total es 1 entonces ningún valor positivo nos sirve puesto que la suma sería infinita y si el valor es 0 tampoco, puesto que entonces la suma nos daría 0 y no 1. De hecho, es imposible modelar este Evento Aleatorio. Esto sólo lo podremos explicar con la axiomática de Teoría de Probabilidades que introduciremos en el último Capítulo, que, además permite tratar Espacios Muestrales infinitos con el mismo tipo de reglas y cálculos que vamos a deducir y utilizar en el caso Discreto. Esta es una de las razones principales por la que se introduce esta formalización.

2.4.2 Identificando Eventos Aleatorios

Si tenemos dos Eventos Aleatorios de naturaleza completamente distinta, pero que se codifican con el mismo Espacio Muestral y para los cuales es igualmente probable obtener cualquier valor, serán para nosotros indistinguibles. Esta es la fundación de nuestro Modelo. Realizar esta asociación produce una reducción drástica de información. Sin embargo, esto resulta ser extremadamente práctico puesto que podemos simular cualquier par de Eventos no distinguibles utilizando al otro.

Resumiendo entonces, para el caso de Espacios Muestrales Discretos, la manera de caracterizarlos es una lista de resultados posibles indicando para cada uno de ellos la posibilidad de ocurrir. Tradicionalmente se introducen argumentos combinatorios para especificar estas posibilidades. En esta monografía vamos a posponer un poco estos argumentos para entregar uno más general y gráfico, que, además, nos permite simular. La idea de simulación será nuestra herramienta principal de intuición y deducción de las reglas de la Teoría de Probabilidades.

2.5 Disco Básico

Basados en la discusión de la sección precedente, resulta claro que es posible representar cualquier Evento Aleatorio Discreto con un disco con una flecha que gira. Este *Disco Aleatorio* estará divido en una cantidad finita de regiones etiquetadas acorde al Espacio Muestral. Una manera simple de construir esta herramienta para poder simular físicamente, es un disco circular tipo ruleta dividido en regiones similares a las obtenidas al cortar una torta de cumpleaños (ver figura 2.2).

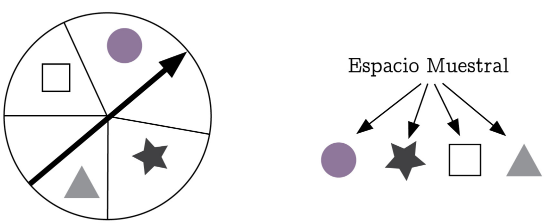


Figura 2.2. Disco Básico

Una realización del Evento Aleatorio a través del Modelo se obtiene girando la flecha en torno al centro del Disco. No se tiene control alguno sobre el destino final de la flecha más que el hecho de saber que se detendrá en algún momento. El resultado del Evento Aleatorio será la etiqueta de la región en que la punta de la flecha se detuvo.

Un Espacio Muestral Discreto también se puede representar como un árbol que tiene ramas que llevan a cada resultado posible (ver figura 2.2). Una realización del Evento Aleatorio, ahora, consiste en seleccionar una de las ramas posibles. Por el momento esto no resulta ser muy interesante, pero más adelante, cuando consideremos más de un Evento Aleatorio, su utilidad quedará de manifiesto.

Volviendo al disco, la comparación de los arcos de las regiones, es, naturalmente, la que cuantifica cuánto más probable es un resultado que otro. En efecto, mientras más grande es el arco de una región respecto de las demás, más probable será obtener ese valor por sobre los otros. De hecho, es bastante intuitivo que esta asociación es lineal. Es decir, si el arco de una región se amplifica al doble, entonces será doblemente más probable que la flecha se detenga en esa región y lo mismo para cualquier otro ponderador.

Es preferible considerar el área de la región como magnitud, que el arco, puesto que resulta más intuitivo y fácil de visualizar. Esto no cambia nada puesto que el arco es proporcional al area de la región.

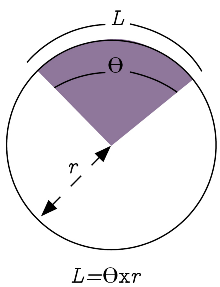


Figura 2.3. Arco de Círculo

Es un ejercicio sencillo probar que el área de una región es proporcional al arco exterior de la misma. El arco de una región dividido por el radio del círculo es el ángulo en radianes de dicho arco (ver figura 2.3). Por otro lado, el área de un sector circular es el producto del ángulo en radianes por el cuadrado del radio del círculo dividido por 2; por tanto estas dos cantidades son proporcionales. En lenguaje matemático si llamamos S al área del sector circular, L al arco y θ el ángulo en radianes del arco tendremos que:

$$S = \frac{1}{2}r^2 \cdot \theta,\ L = r \cdot \theta \Rightarrow S = \frac{1}{2}r^2 \cdot \frac{L}{r} = \frac{r}{2}P. \tag{2.3}$$

Siendo que sólo se consideran proporciones de áreas o arcos en el Disco, amplificar de manera geométrica el Disco no cambia los resultados, por lo que es natural fijar una escala. La escala usual para probabilidades es considerar la probabilidad de que la flecha se detenga en una región como el cociente entre el área que ocupa la etiqueta y el área total del disco. Según esta convención el valor menos probable será 0 y el más probable 1.

La máxima información sobre un Evento Aleatorio, o lo que es lo mismo, la mínima incertidumbre, es un disco con una sola etiqueta. Es cierto que parece un poco absurdo considerar este caso como aleatorio, pero es absolutamente necesario hacerlo, como es necesario considerar el conjunto vacío en la Teoría de Conjuntos.

Nota: Por vacuidad cualquier etiqueta que no esté en el Disco Aleatorio tiene un valor probable de 0 (estas etiquetas son resultados imposibles).

2.6 Árbol y Tabla de un Evento Aleatorio Discreto

Si bien es cierto, el Disco Aleatorio es una representación muy intuitiva que permite simular una o más realizaciones de un Evento Aleatorio, no es muy útil llegado el momento de realizar cálculos numéricos.

En la representación de árbol del Espacio Muestral, es posible agregar la probabilidad de cada resultado a cada uno los arcos que indica cuán probable es tomar ese camino en una realización. Con esto sacrificamos la parte gráfica de la simulación obteniendo a cambio del valor numérico exacto de cada resultado posible.

Eliminar por completo la simulación corresponde a construir una tabla de dos filas. En la primera fila se colocan las etiquetas del disco y en la segunda la probabilidad de obtener cada etiqueta (ver figura 2.4).

Estas tres representaciones son absolutamente equivalentes respecto a la información que entregan, pero su utilidad es por completo diferente. De hecho pasar de una a otra es bastante inmediato.

En el caso del Disco a Tabla o Árbol es necesario calcular áreas para convertirlas a números. En dirección contraria basta utilizar un trazo recto cualquiera dividido en tantos trazos como etiquetas, manteniendo entre los trazos la proporción dada por la tabla. Una vez realizado esto se unen los extremos del trazo para formar un círculo y se marcan las regiones obtenidas con las etiquetas correspondientes.

Nota pedagógica: Existe un sinfín de actividades pedagógicas que se pueden realizar utilizando estas tres representaciones y su relación. Por ejemplo, la construcción de un Disco Aleatorio, ya sea en papel o por computador, a partir de la información de la tabla y viceversa.

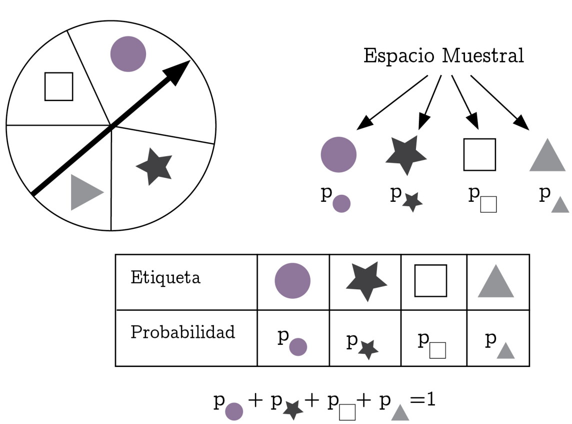


Figura 2.4. Tabla, Árbol y Disco de un Evento Aleatorio

2.7 Ejemplos Clásicos de Juegos de Azar

Una manera de poner en práctica lo aprendido y recorrer otra vez los caminos de la historia de las probabilidades, es conectarlo con elementos como monedas, dados y mazos de cartas.

2.7.1 Monedas

Apostar al lanzamiento de una moneda es un juego muy antiguo. Los romanos lo llamaban "navia aut caput" (barco o cabeza), dado que las monedas tenían un barco por un lado y la efigie del emperador en la otra. Es habitual que la persona que escoge no es la persona que lanza la moneda para evitar posibles trampas. Este elemento es muy utilizado como herramienta para zanjar disputas o para tomar decisiones de manera imparcial hasta el día de hoy. En los deportes es común que el árbitro lance una moneda para decidir quien comienza con el balón y/o escoge el lado de la cancha.

Al jugar el clásico juego de cara o sello estamos considerando un Espacio Muestral con dos sucesos posibles. El Disco Aleatorio tendrá dos regiones una marcada con "CARA" y otra con "SELLO".

En una versión justa de este juego la moneda está *equilibrada* y entonces las dos regiones tendrán áreas iguales. Si la suma total es 1 entonces la probabilidad de cada uno de estos dos valores es $\frac{1}{2}$ y para cualquier otro valor será 0. Si la moneda no está equilibrada una región tendrá probabilidad $0 \leq p \leq 1$ y la otra $1 - p$.

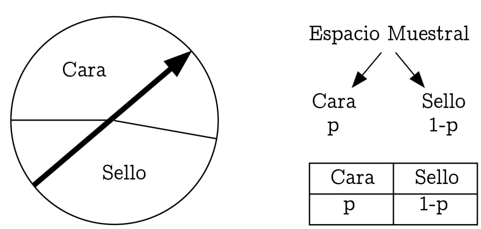


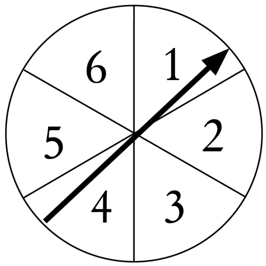
FIGURA 2.5. Disco y Moneda

2.7.2 Dados

Un dado clásico tiene 6 caras numeradas 1,2,3,4,5 y 6 que se asumen igualmente probables por la simetría del dado. Esto equivale a un Disco Aleatorio con estas etiquetas y con todas sus regiones iguales (ver figura 2.6).

2.7.3 Cartas

Los mazos de cartas son otro ejemplo de azar. En el caso de nuestro país estamos hablando de la baraja internacional de 52 cartas. Cada carta tiene una de 4 pintas



1	2	3	4	5	6
1/6	1/6	1/6	1/6	1/6	1/6

Figura 2.6. Tabla y Disco Aleatorio de un dado

(CORAZONES $\heartsuit$, DIAMANTES $\diamondsuit$, PICAS $\spadesuit$ y TRÉBOLES $\clubsuit$) y 13 símbolos (A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K).

Extraer una carta del mazo tiene asociado un Disco Aleatorio con 52 divisiones cada una con probabilidad $\frac{1}{52}$.

2.8 Preguntas y Respuestas en un Experimento

Cualquiera de las tres representaciones de un Evento Aleatorio permite, para cualquier valor posible de x tanto fuera como dentro del Espacio Muestral, responder inmediatamente a la pregunta ¿cuál es la probabilidad de que el Evento Aleatorio tome el valor x?. Aunque en el caso del disco debamos medir el valor numérico del área.

De hecho, estas son las únicas preguntas que se pueden responder sin cálculo aritmético alguno. Con la información del Disco podemos responder otras preguntas más complejas mediante ciertos cálculos y habrá otras que no podremos responder. A un mismo Evento Aleatorio se le pueden asociar diferentes Espacios Muestrales y por tanto Discos Aleatorios. La información cuantitativa que se incorpore a cualquiera de estas representaciones debe ser consistente y compatible. Muchas veces seleccionar el Espacio Muestral correcto es crucial para responder adecuadamente un problema y evitar cometer errores.

2.8.1 Preguntas y Reetiquetado

Si las preguntas a responder están decididas, son estas las que definen si el Espacio Muestral es adecuado o no. El criterio para decidir es bastante simple. Basta una región del disco donde no se determina la respuesta a una pregunta, para que el Espacio Muestral no sea adecuado. Por ejemplo, si al lanzar un dado de seis caras con el Evento Aleatorio asociado a los resultados posibles del número de la cara superior, me pregunto si cayó en la mesa o afuera!. En este caso, en todas las regiones es posible que el dado esté en la mesa o afuera. La "recíproca" a esta pregunta es determinar las preguntas que se pueden responder, conociendo el Espacio Muestral. La manera más directa y tradicional de seleccionar un Espacio Muestral para un Evento Aleatorio es justamente fijar las preguntas que se quieren responder.

Observemos que una vez definido el Espacio Muestral no parece fácil dar el listado de todas las preguntas que se pueden responder con este. Sin embargo, por nuestra observación, podemos dar una lista de todas las respuestas posibles que podemos entregar. En efecto, numéricamente deben ser sumas y divisiones de los valores numéricos de las áreas de los sectores circulares presentes en el disco. La suma proviene de agrupar las áreas y la división de renormalizar las áreas al eliminar alguna de ellas.

Dado un Disco, un *reetiquetado* (ver figura 2.7) consiste en cambiar las etiquetas del disco agrupando las áreas que quedan con la misma etiqueta. Agrupar es necesario puesto que definimos los Discos de modo que cada región tenga una etiqueta diferente. El área de la agrupación de las regiones es obviamente la suma de las áreas que la forman.

Nota: Esta propiedad sobre la agrupación de áreas se conoce como la *Propiedad Aditiva* de las probabilidades.

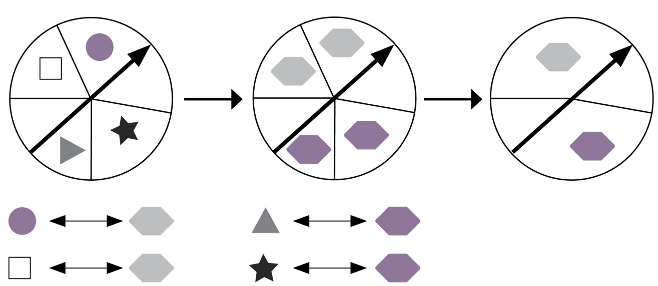


FIGURA 2.7. Reetiquetado de un Disco

La operación sobre la representación de Tabla es igual de simple (ver figura 2.8).

Otra manera habitual de incorporar información a un Evento Aleatorio, corresponde a entregar información parcial que elimina la posibilidad de que algunos de los resultados del Espacio Muestral hayan ocurrido. Naturalmente, los resultados restantes mantienen la proporción de ocurrir que tenían originalmente pero ahora son todos los resultados posibles y por tanto deben sumar 1. Esto corresponde a eliminar regiones del Disco y renormalizar las restantes para armar un nuevo Disco.

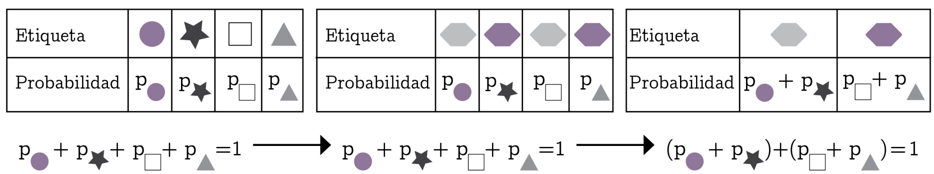


Figura 2.8. Operación reetiquetado en tablas

Una *simplificación* de un Disco (ver figura 2.9) consiste en eliminar etiquetas llenando el espacio del disco entre las demás regiones respetando la proporción existente entre las regiones sobrevivientes.

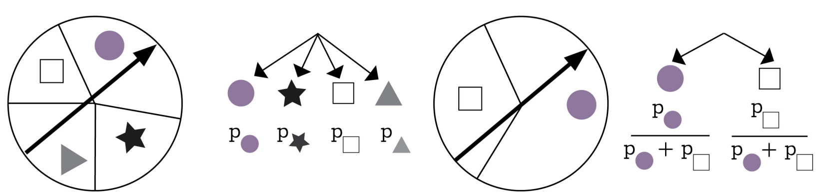


Figura 2.9. Simplificación de un disco

La operación de simplificación sobre la tabla corresponde simplemente a una renormalización de las áreas de las regiones restantes, que se realiza dividiendo cada área restante por la suma de las áreas de ellas (ver figura 2.10).

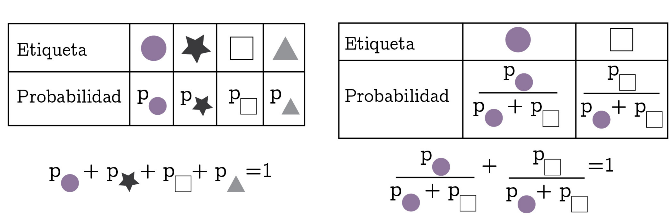


Figura 2.10. Operación simplificación en tablas

Nota: La operación de simplificación de un Disco corresponde a lo que se conoce como *Probabilidad Condicional*.

Toda respuesta a una pregunta compatible con el Modelo se obtiene combinando estas dos operaciones. Veremos a continuación que, con estas dos operaciones, es posible resolver una gran parte de los problemas de la Teoría de Probabilidades, sin tener que introducir formalmente ni Combinatoria ni la axiomática de la Teoría de Conjuntos. En las secciones que siguen, vamos a seguir desarrollando esta metodología utilizando las herramientas que acabamos de definir entrenando además la habilidad de seleccionar el Espacio Muestral adecuado a cada situación.

Nota pedagógica: Las operaciones de reetiquetado y simplificación sobre discos, como actividad pedagógica, son interesantes a realizar por los alumnos ya sea con computador o manualmente. Permite al alumno manipular y relacionar probabilidades con otras áreas como Geometría y Artes Manuales. Además, tiene aplicaciones relacionadas con propiedades como asociatividad, conmutatividad y distributividad de las operaciones de suma y producto.

2.9 El Ejemplo de la Extracción de Bolitas

Para comenzar, veamos un ejemplo concreto. Consideremos una urna con 15 bolitas; 1 ploma, 2 negras, 3 blancas, 4 moradas y 5 grises. Nuestro Evento consiste en sacar una bolita de la urna sin saber cuál es.

Las preguntas que vamos a responder son ¿Cual es la probabilidad de que la bolita sea ploma, negra, blanca, morada o gris? Un primer Disco esta formado por 15 regiones que representan a cada una de las 15 bolitas en una extracción. Vamos a numerar las bolitas ordenadas por el color, es decir, la 1 es ploma, 2 y 3 negras, 4, 5 y 6 blancas, etc...

Sin disponer de más información todas las regiones tienen la misma área, $\frac{1}{15}$ del área total para ser precisos. No vamos a dibujar este disco por razones obvias (15 regiones!!) pero la tabla será:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/	1/15

Veamos cómo responder las preguntas que motivaron este Modelo. Sabemos exactamente el color de cada bolita y por tanto cuáles son las regiones que corresponden a obtener cada uno de los colores posibles. Aplicando el reetiquetado de cada bolita por su color obtenemos un Disco con 5 regiones etiquetadas con los colores posibles de las bolitas 2.11. Utilizando la tabla original podemos construir la tabla de este nuevo disco fácilmente:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/15	1/	1/15

PLOMA	NEGRA	BLANCA	MORADA	GRIS
1/15	2/15	3/15	4/15	5/15

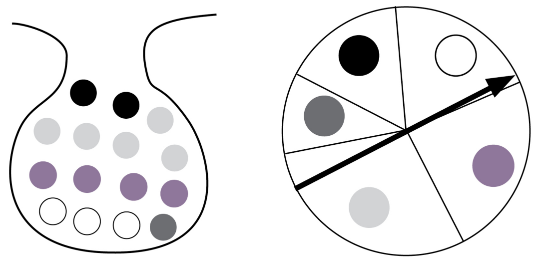


FIGURA 2.11. Disco Aleatorio de colores posibles

Concluimos entonces que la probabilidad de obtener una bolita ploma, negra, blanca, morada o gris es respectivamente $\frac{1}{15}$, $\frac{2}{15}$, $\frac{3}{15}$, $\frac{4}{15}$ y $\frac{5}{15}$.

Es cierto que podríamos haber comenzado con este disco de 5 regiones desde un principio. Pero en ese caso sería necesario dar el argumento lógico que nos permite deducir que no estamos en un caso equiprobable (si lo piensan un poco eso es en el fondo lo que acabamos de hacer).

Si inicialmente representamos el Disco con 5 regiones sin mayor información, se asumiría que cada uno de los colores es tan probable de ocurrir como cualquier otro. Es decir, la siguiente tabla:

PLOMA	NEGRA	BLANCA	MORADA	GRIS
1/5	1/5	1/5	1/5	1/5

Este Disco Aleatorio entrega respuestas diferentes al que obtuvimos anteriormente porque considera otra información. Parece un poco rebuscado y excesivo todo este procedimiento de cálculo, cuando era más fácil utilizar el consabido cálculo de *favorables sobre posibles* para el primer Evento Aleatorio y sumar.

Sin embargo, el procedimiento que acabamos de hacer nos permite resolver el problema, de la misma manera, en situaciones que el cálculo de favorables sobre posibles da respuestas erradas.

Por ejemplo: ¿Cómo procedemos si es a veces más probable obtener una bolita negra que una ploma, b veces más probable obtener una bolita blanca que una negra, c veces más probable obtener una morada que una blanca y d veces más probable obtener una gris que una morada?

Si llamamos P el área de una de las bolitas ploma, N el área de una de las bolitas negras, B el área de una de las bolitas blancas, M el área de una de las bolitas moradas y G el área de una de las bolitas grises. Entonces la tabla de nuestro primer Disco Aleatorio ahora es:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	N	N	B	B	B	M	M	M	M	G	G	G	G	G

Sabemos que $P+2N+3B+4M+5G=1$ y si incorporamos la nueva información tenemos que $N=aP$, $B=bN$, $M=cB$, $G=dM$. Lo que con un poco de álgebra básica nos permite dejar todas las áreas en función del área P. En efecto $N=aP$,$B=abP$,$M=abcP$ y $G=abcdP$ y por lo tanto:

$$P=\frac{1}{1+a+ab+abc+abcd}:=\frac{1}{r}$$

La suma bajo la fracción la nombramos como r. Entonces:

PLOMO	NEGRA	BLANCA	MORADA	GRIS
$1/r$	$2a/r$	$3ab/r$	$4abc/r$	$5abcd/r$

2.9.1 Ejercicios Propuestos

1. Muestre que para obtener igual probabilidad de extraer una bola de cada color es necesario y suficiente que $a=\frac{1}{2}$, $b=\frac{2}{3}$, $c=\frac{3}{4}$ y $d=\frac{4}{5}$. Justifique en palabras este resultado.
2. Un defensor acérrimo de la regla de favorables sobre posibles insiste en que es posible resolver el problema con dicha regla aún en el caso no equiprobable. Para esto sugiere que lo único necesario es agregar más cantidad de bolitas para cada color, asumiendo que son todas igualmente probables. Determine si tiene razón y encuentre la cantidad de bolitas que es necesario agregar en ese caso.

2.10 Combinando Eventos Aleatorios

De aquí en adelante sólo vamos a considerar el caso de Espacios Muestrales Discretos por lo que vamos omitir el vocablo Discreto.

Si bien es cierto que todo Evento Aleatorio se puede representar con un Disco Aleatorio solamente, en esta representación toda la incertidumbre está en un sólo giro de la aguja. En muchas ocasiones resulta más útil y/o natural descomponerlo como una secuencia de Eventos Aleatorios.

¿Cómo modelamos más de un Evento Aleatorio? La respuesta obvia parece ser considerar un disco por cada Evento Aleatorio. La realización conjunta de los dos Eventos en secuencia corresponde a girar las agujas de dos discos y el resultado obtenido es un par de etiquetas de ambos discos en el orden de secuencia.

Esto pareciera ser un Modelo adecuado y una generalización correcta de nuestro enfoque para un Evento Aleatorio. Sin embargo, si invertimos el orden de realización de los Eventos veremos que obtenemos la misma simulación (excepto en el orden de las etiquetas por supuesto). Esto en particular indica que con este Modelo los resultados de un Evento no pueden influir en el otro (ver figura 2.12).

La aplicación en secuencia, naturalmente, nos indica que hay que formar un Disco para el primer Evento Aleatorio. Al interior de cada región se ha determinado el

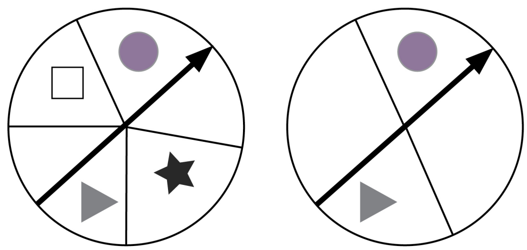


Figura 2.12. Dos Discos Aleatorios en Paralelo

resultado del primer Evento, pero no se sabe nada del segundo y por tanto la falta de información al interior de cada región corresponde a otro Disco asociado al segundo Evento. Lo interesante es que en cada región del primer disco, habrá un disco diferente que cuantifica la influencia del primer Evento sobre el segundo. En el caso en que los discos al interior de las regiones del primer disco sean todos iguales nos encontramos con una representación equivalente a la que habíamos escogido puesto que podríamos “factorizar” ese disco interior y dejarlo al exterior en secuencia. Veremos más adelante a qué caso particular de dos Eventos Aleatorios corresponde esta situación.

Entonces, de manera general, dos Eventos Aleatorios en secuencia quedan representados por un Disco Aleatorio habitual que en cada una de sus regiones tiene otro Disco Aleatorio. Una realización de este *Disco Compuesto* corresponde a girar dos flechas en secuencia, la flecha del primer Disco y luego la flecha del Disco que corresponde a la región donde la primera flecha se detuvo.

Al representar un Evento Aleatorio Compuesto como árbol, tenemos dos niveles puesto que la realización es en dos etapas. En el primer nivel tenemos un camino por cada etiqueta del primer Disco y saliendo de cada nodo del primer nivel, tenemos un camino por cada etiqueta del segundo Disco asociado a ese nodo (ver figura 2.13).

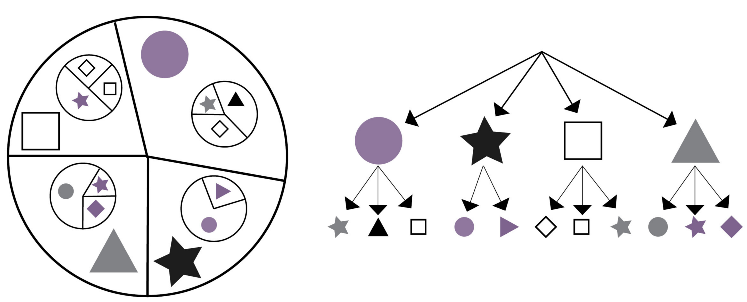


Figura 2.13. Un Evento Aleatorio Compuesto

Resulta directo y bastante intuitivo obtener la probabilidad de un resultado conjunto. La probabilidad de obtener el primer resultado es el área de la región asociada, pero si además se va a obtener un segundo resultado, esta área se divide proporcionalmente al segundo Disco en la región. En otras palabras la probabilidad de obtener un par de resultados es simplemente el producto de las probabilidades de cada una de las regiones del Disco que entregan dicho resultado.

Nota: Esto se conoce como la *Propiedad Multiplicativa de la Probabilidad Condicional.*

Al considerar los Eventos como un árbol, los resultados posibles son todos los pares de caminos y la probabilidad del camino compuesto es el producto de los caminos que lo forman. Los Discos interiores de un Evento Aleatorio Compuesto, cuantifican la información condicionada al resultado de la región para el segundo Evento. Es decir son *Discos Condicionales* del segundo Evento respecto al primero.

La primera representación como tabla de dos Eventos Aleatorios en secuencia la llamaremos *Tabla Condicional del Disco Compuesto* y corresponde a agregar una fila más a la tabla del primer Evento, en que los valores de dicha fila, son las tablas asociadas a los discos al interior de cada región del disco (ver figura 2.14). como en el caso del árbol.

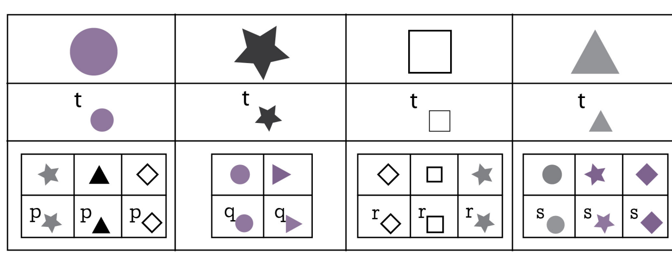


FIGURA 2.14. Tabla Condicional de dos Eventos Aleatorios en secuencia

2.10.1 Eventos Aleatorios Compuestos de más de dos Eventos

El mismo razonamiento que utilizamos para simular y representar dos Eventos Aleatorios en secuencia, se puede utilizar en general para una secuencia finita de los mismos. Esto resulta más fácil de visualizar y comprender como árbol, puesto que cada nivel del árbol representa un Evento y por cada nodo sale una rama al evento siguiente. La concatenación de las etiquetas obtenidas es el resultado del Evento Compuesto y la probabilidad asociada es, naturalmente, el producto de las probabilidades de cada rama. En el caso de un disco esto resulta complejo y engorroso de representar, dado que son discos dentro de discos. Afortunadamente no es necesario considerar representaciones de más de dos Discos. Veremos que cualquier Evento Aleatorio, que se representa por una serie finita de Eventos Aleatorios, se puede llevar recursivamente a sólo dos y de hecho la representación en dos resulta ser la adecuada y necesaria para realizar cálculos. Sin embargo, la estructura de árbol con más de dos niveles será de gran utilidad y muy aprovechada en el Capítulo de Combinatoria.

2.10.2 Independencia de Eventos Aleatorios Discretos

La realización en paralelo, que planteamos inicialmente, será consistente solo cuando los discos al interior de las regiones del primer disco sean todos iguales (ver figura 2.15).

Como vimos, cuando esto ocurre tenemos que el orden de realización entre los dos eventos no cambia la cuantificación de obtener el resultado conjunto. Este es el caso de *Eventos Aleatorios Independientes*. En la representación de árbol esto ocurre cuando en el segundo nivel tenemos copias iguales en cada nodo.

Cuando todos los Discos Condicionales coinciden, lo hacen con el disco del Segundo Evento Aleatorio realizado como en el orden inverso, es decir, de manera *independiente* del primero que de alguna manera justifica este nombre. Por la regla multiplicativa obtenemos que, la probabilidad de obtener un par de resultados de dos Eventos independientes, es simplemente el producto de sus probabilidades. De hecho esta propiedad resulta ser además de necesaria, suficiente para que dos Eventos sean Independientes.

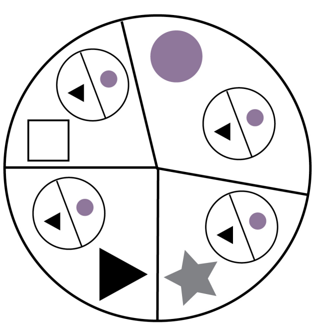


FIGURA 2.15. Dos Eventos Aleatorios Independientes en secuencia

2.11 Fusionando Eventos Aleatorios Compuestos

En la sección anterior vimos como representar dos Eventos Aleatorios utilizando un par de Discos e introdujimos una representación de árbol asociada y una tabla. Utilizamos fuertemente la idea de considerar los resultados posibles de dos Discos como la concatenación de sus etiquetas. Es natural entonces, considerar estos resultados concatenados como uno sólo, obteniendo un solo Disco asociado al Disco Compuesto.

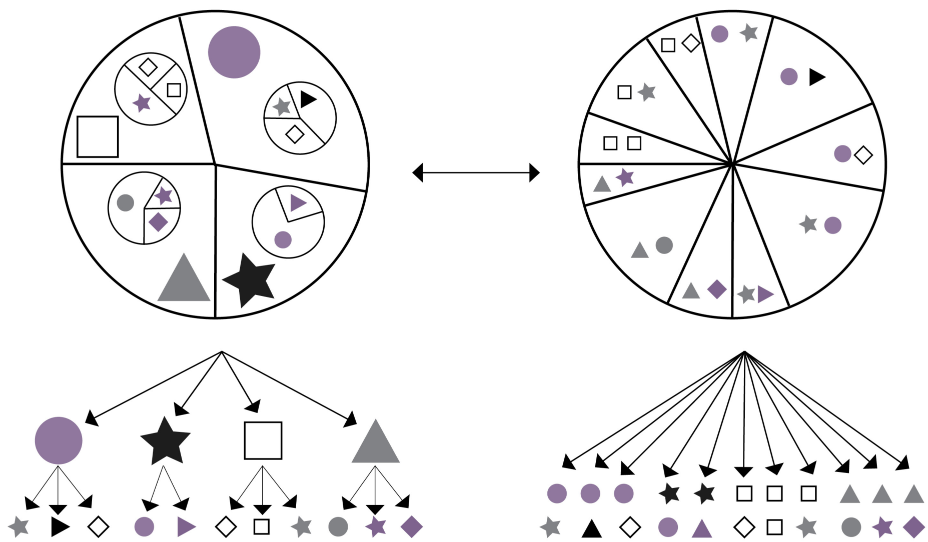


FIGURA 2.16. Fusionando un Evento Aleatorio Compuesto

En este nuevo Disco el área de región se obtiene por simple multiplicación. Cada sector del primer Disco se divide en la cantidad de regiones del disco al interior de cada región, respetando la proporción de área entre ellos (ver figura 2.16). Es posible escribir la tabla de dos filas que estábamos utilizando para este tipo de Discos que ahora llamaremos *Tabla Marginal del Disco Compuesto*. Es sorprendente, que aunque esta tabla tiene una fila menos que la Tabla Condicional, resulta que no pierde información alguna.

Un primer paso para desacoplar el orden preestablecido entre los Eventos, es considerarlo como una tabla de doble entrada en la que las etiquetas del primer Disco están en la horizontal y las etiquetas del segundo Disco están en la vertical. Ahora, invertir el orden en que se realizan los Eventos es, simplemente cambiar las filas por las columnas. Esta simple operación en una tabla, se conoce como *transposición*. El valor en una casilla cualquiera de esta tabla corresponde a la probabilidad de obtener el par de resultados que indica la casilla. Llamaremos a esto *Tabla Conjunta del Disco Compuesto*.

Esta tabla bidimensional, en la gran parte de los casos, genera un Espacio Muestral más grande que los resultados posibles del Evento Aleatorio Compuesto que se

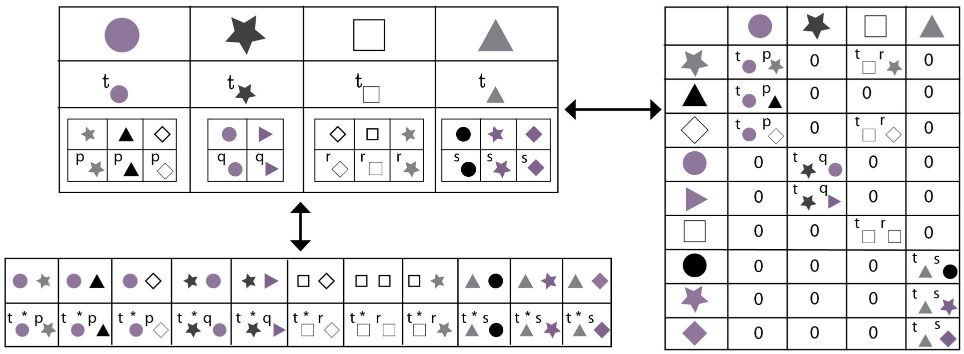


Figura 2.17. Tabla Condicional, Conjunta y Marginal de dos Eventos Aleatorios en secuencia

agregan con probabilidad nula. Es evidente que la información de esta tabla bidimensional, es la misma que la tabla del Disco Aleatorio fusionado, puesto que tiene los mismos valores de casilla agregando posiblemente algunos ceros (ver figura 2.17).

Hasta el momento, hemos visto cómo obtener la Tabla Conjunta y la Marginal a partir de la Tabla Condicional. Si no se pierde información en este proceso debiéramos ser capaces de revertirlo y recuperar la Condicional de cualquiera de estas dos nuevas tablas.

Si inicialmente tenemos la Tabla Conjunta, veamos cómo construir la Tabla Condicional. Los resultados del primer evento corresponden a las etiquetas de la fila, por lo que ésta se preserva en la Tabla Conjunta. Para obtener las probabilidades de cualquier resultado del primer Evento Aleatorio, basta sumar los valores de la Tabla Conjunta que tienen esta etiqueta y que son todos resultados distintos (ésta era la propiedad aditiva); es decir sumar por columnas.

Resumiendo, la probabilidad de obtener cada resultado del primer evento corresponde a la suma de los valores de la columna correspondiente, mientras que la segunda fila de la Tabla Condicional es entonces la suma de todas las columnas correspondientes en la Tabla Marginal.

Nota: Esto se conoce como la *Fórmula de Probabilidades Totales.*

Las etiquetas del Disco Condicional, en cada columna, corresponden a las etiquetas verticales que en esa columna tengan valores no nulos de probabilidad. En esta etapa se produce una transposición, porque estos valores que están ordenados en columna pasan a estar en la primera fila del Disco Condicional.

La suma de los valores no nulos de cada columna entrega la probabilidad de obtener el resultado de esa columna. Si todos estos valores se dividen por esta suma se obtienen valores que respetan la proporción pero que ahora suman 1. Esto corresponde a los valores del Disco Condicional para cada columna en la Tabla Condicional.

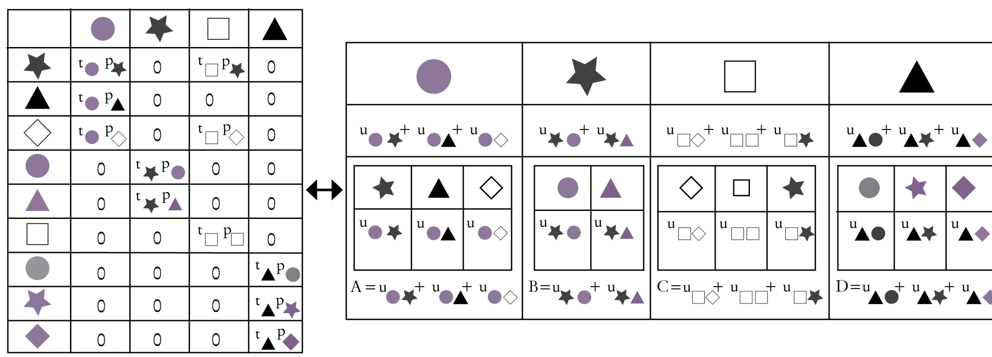


Figura 2.18. Pasaje de Tabla Conjunta a Tabla Condicional

2.11.1 El Disco de Bayes

Dos Eventos Aleatorios en secuencia tienen un orden establecido y es posible que en los Eventos que los definen, sean física o lógicamente imposibles de invertir.

Curiosamente, como ya habíamos mencionado antes, al representar los dos Eventos en Tabla Conjunta, invertir el orden de realización solo requiere de transponer la tabla (ver figura 2.19). La Tabla Marginal asociada a esta tabla transpuesta la llamaremos *Tabla Marginal de Bayes del Disco Compuesto*.

El Disco Aleatorio, asociado a la Tabla Marginal de Bayes (*Disco de Bayes*), corresponde a realizar los dos Eventos Aleatorios en el orden inverso de manera consistente. Con esto, queremos decir que todos los cálculos de probabilidades se mantienen.

Los Discos Condicionales del Disco de Bayes nos entregan las probabilidades de obtener un resultado del primer Evento (en el orden original) bajo el supuesto que un resultado del segundo Evento ocurrió.

Nota: Esto se conoce como la *Regla de Bayes*.

Un procedimiento más geométrico (y equivalente) al descrito anteriormente, se obtiene al invertir el orden de realización simplemente intercambiando el orden de las etiquetas, para luego reconstruir la Tabla Condicional.

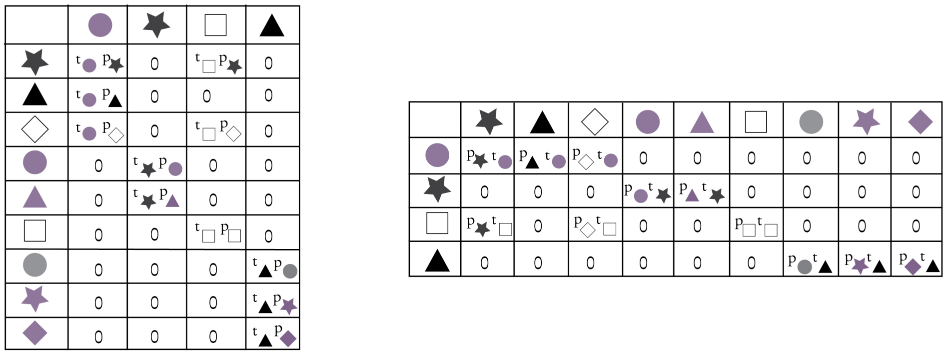


Figura 2.19. Inversión del orden de realización en Tabla Conjunta

A nivel de Discos, el procedimiento consiste en fusionar el Disco Compuesto, en su versión simple, para luego invertir las etiquetas y finalmente formar el Disco Compuesto respecto a esta nueva primera etiqueta (ver figura 2.20).

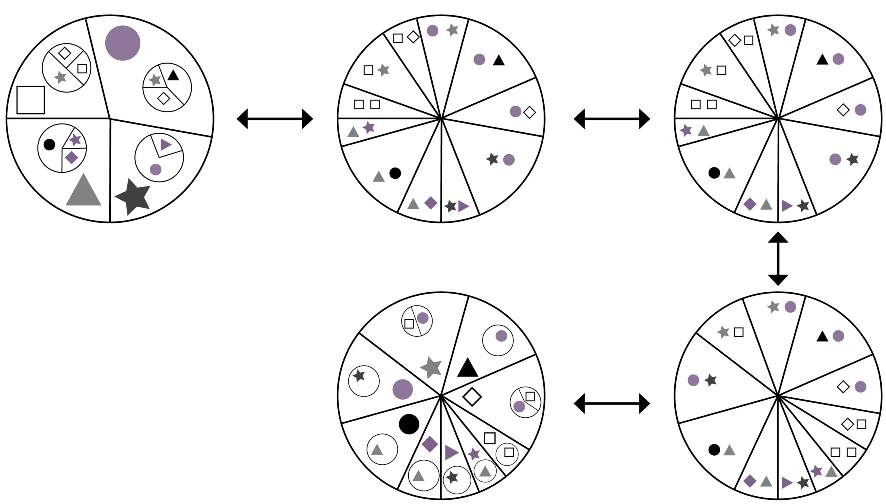


Figura 2.20. Inversión del orden de realización en Disco Aleatorio

Nota: Muchas actividades pedagógicas interesantes respecto a Discos Compuestos se pueden realizar con alumnos al dar un Disco a cada alumno y simular los Eventos Aleatorios dando el turno a cada alumno cuando corresponda.

2.12 Ejemplos Clásicos de Juegos de Azar II

Una vez más, para poner en práctica las definiciones, vamos a utilizar los elementos lúdicos clásicos que ya habíamos introducido, considerando ahora, dos o más realizaciones.

2.12.1 Monedas

Lanzar una moneda dos veces corresponde a dos Eventos Aleatorios en secuencia. Esto se representa como un Disco Aleatorio con dos regiones etiquetadas "CARA" y "SELLO"; y áreas p y $1-p$, respectivamente y una copia de ese mismo disco en cada una de las dos regiones. Al fusionar el Disco, obtenemos un disco con 4 regiones con etiquetas: "CARA CARA", "CARA SELLO", SELLO CARA" y "SELLO SELLO" (ver figura 2.21).

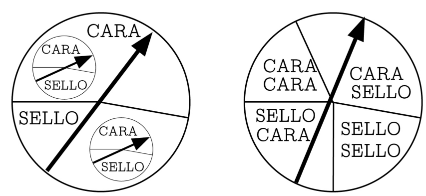


Figura 2.21. Disco Aleatorio del lanzamiento de dos monedas

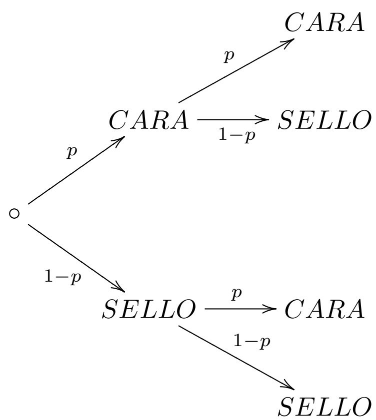
Entonces la etiqueta "CARA", en el Disco Compuesto de área p, se subdivide en proporción p y $1-p$ lo que da p^2 y $p(1-p)$ para las etiquetas "CARA CARA" y "CARA SELLO" respectivamente y de manera análoga obtenemos áreas $(1-p)^2$ y $p(1-p)$ para "SELLO SELLO" y "SELLO CARA". Entonces las Tablas Condicional, Marginal y Conjunta son:

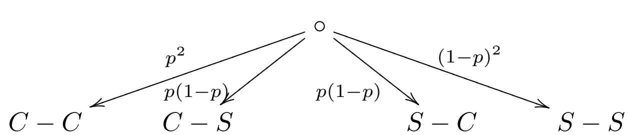
<table>
<tr><td colspan="2">CARA</td><td colspan="2">SELLO</td></tr>
<tr><td colspan="2">p</td><td colspan="2">$1-p$</td></tr>
<tr><td>CARA</td><td>SELLO</td><td>CARA</td><td>SELLO</td></tr>
<tr><td>p</td><td>$1-p$</td><td>p</td><td>$1-p$</td></tr>
</table>

CARA CARA	CARA SELLO	SELLO CARA	SELLO SELLO
p^2	$p(1-p)$	$(1-p)p$	$(1-p)^2$

	CARA	SELLO
CARA	p^2	$p(1-p)$
SELLO	$(1-p)p$	$(1-p)^2$

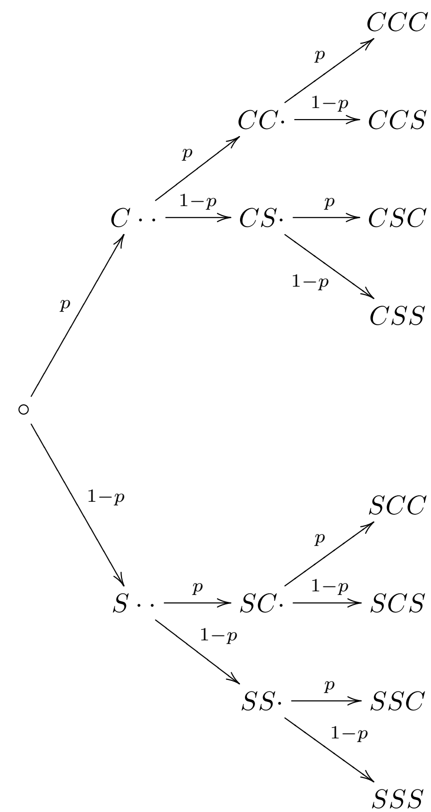
Los pasos equivalentes en el diagrama de árbol son:





En este ejemplo se observa que dos realizaciones del mismo Evento Aleatorio en estas condiciones resultan independientes. Esto se puede ver desde un principio, dado que los Discos Condicionales son idénticos en ambas regiones.

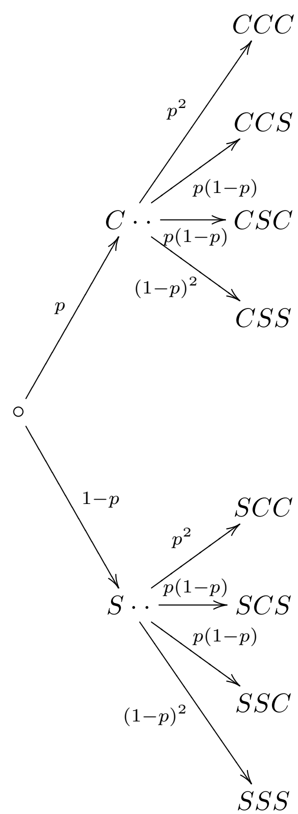
Como ya mencionamos, es posible considerar el Evento Aleatorio del lanzamiento de una moneda 3 veces sólo con un Disco Compuesto. Si notamos los resultados posibles como $C\cdot\cdot$,$S\cdot\cdot$,$\cdot C\cdot$,$\cdot S\cdot$,$\cdot\cdot C$ y $\cdot\cdot S$. Es decir $C\cdot\cdot$ es cara en en primer lanzamiento y $\cdot SC$ es sello en el segundo y cara en el tercero y así sucesivamente. El árbol de todas las posibilidades es:



En formato de Tabla Condicional nos da:

<table>
<tr><td colspan="4">$C\cdot\cdot$</td><td colspan="4">$S\cdot\cdot$</td></tr>
<tr><td colspan="4">p</td><td colspan="4">(1-p)</td></tr>
<tr><td colspan="2">$\cdot C\cdot$</td><td colspan="2">$\cdot S\cdot$</td><td colspan="2">$\cdot C\cdot$</td><td colspan="2">$\cdot S\cdot$</td></tr>
<tr><td colspan="2">p</td><td colspan="2">$(1-p)$</td><td colspan="2">p</td><td colspan="2">$(1-p)$</td></tr>
<tr><td>$\cdot\cdot C$</td><td>$\cdot\cdot S$</td><td>$\cdot\cdot C$</td><td>$\cdot\cdot S$</td><td>$\cdot\cdot C$</td><td>$\cdot\cdot S$</td><td>$\cdot\cdot C$</td><td>$\cdot\cdot S$</td></tr>
<tr><td>p</td><td>(1-p)</td><td>p</td><td>(1-p)</td><td>p</td><td>(1-p)</td><td>p</td><td>(1-p)</td></tr>
</table>

Al colapsar los dos últimos niveles del árbol obtenemos:



Su equivalente como Tabla Marginal nos da:

$C\cdot\cdot$				$S\cdot\cdot$			
p				(1-p)			
$\cdot CC$	$\cdot CS$	$\cdot SC$	$\cdot SS$	$\cdot CC$	$\cdot CS$	$\cdot SC$	$\cdot SS$
p^2	$p(1-p)$	$p(1-p)$	$(1-p)^2$	p^2	$p(1-p)$	$p(1-p)$	$(1-p)^2$

Como se puede apreciar de este ejemplo, la representación de árbol es extremadamente engorrosa de dibujar en comparación con la Tabla Marginal que es más compacta y ordenada. Además, ambas representan el mismo enfoque.

2.12.2 Dados

El juego clásico de dados consiste en lanzar dos dados equilibrados de seis caras. El Disco Aleatorio Compuesto asociado es un disco con 6 regiones de igual area con las etiquetas de un dado y un Disco igual al interior de cada región (ver figura 2.22). El Disco fusionado tiene 36 areas iguales, etiquetadas de la forma ij con i y j con valores entre 1 y 6.

La Tabla Conjunta sera:

	1	2	3	4	5	6
1	1/6	1/6	1/6	1/6	1/6	1/6
2	1/6	1/6	1/6	1/6	1/6	1/6
3	1/6	1/6	1/6	1/6	1/6	1/6
4	1/6	1/6	1/6	1/6	1/6	1/6
5	1/6	1/6	1/6	1/6	1/6	1/6
6	1/6	1/6	1/6	1/6	1/6	1/6

Un juego clásico con dos dados es apostar al valor de la suma de éstos, ¿cuan probable es ganar apostando a cada uno de los valores de la suma?

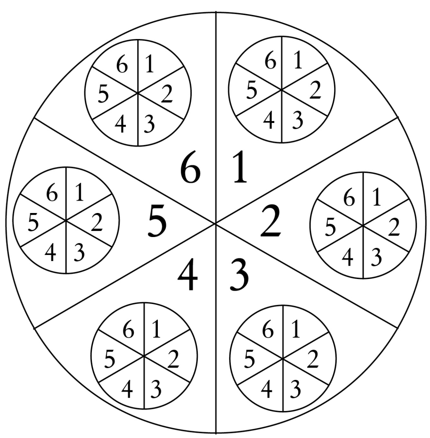


Figura 2.22. Metadisco de dos dados

Los valores posibles para la suma van desde 2 hasta 12 lo que nos da 11 resultados posibles. Usando la estructura de la Tabla Conjunta podemos ver el valor de la suma para cada uno de estos resultados:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Al agrupar los pares de resultados que dan la misma suma obtenemos la Tabla Marginal:

2	3	4	5	6	7	8	9	10	11	12
1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36.

Esto nos muestra que lo más conveniente es apostar al 7.

Nota pedagógica: Wolfgang Amadeus Mozart (1756-1791) compuso la obra Musikalisches Würfelspiel de una manera tan ingeniosa, que no solo lo llevó a componer una pieza para piano, sino un generador de valses. Mozart inventó un juego en el que se lanzan dos dados y que permite componer un vals sin tener conocimientos de música!!. Esto se conoce como el *Juego de los Dados de Mozart.*

2.12.3 Cartas

El experimento de sacar una carta al azar de un mazo de 52 cartas resulta un poco engorroso si se considera como un solo Disco con tantas divisiones y no se aprovecha la estructura de la baraja. Este es un perfecto ejemplo de cómo resulta a veces más interesante representar un Evento de manera Compuesta, aún si no está planteado físicamente de este modo.

Seleccionar una carta del mazo se puede plantear como escoger el valor y luego la pinta. Modelamos esto con un Disco con 13 divisiones iguales etiquetadas como A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q y K con un mismo Disco al interior de cada región con 4 regiones iguales, etiquetadas $\heartsuit$, $\diamondsuit$, $\spadesuit$ y $\clubsuit$.

La Tabla Conjunta de este Evento es:

	A	2	3	4	...	J	Q	K
$\heartsuit$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	...	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$
$\diamondsuit$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	...	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$
$\spadesuit$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	...	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$
$\clubsuit$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	...	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$	$\frac{1}{13}\cdot\frac{1}{4}$

Capítulo 3: Los Problemas

A continuación vamos a resolver algunos *Problemas Emblemáticos* de la Teoría para poner a prueba y ejemplificar el uso de la metodología de cálculo que hemos desarrollado. Veremos cómo se aclaran algunas paradojas, la mayor parte de ellas, mostrando que sólo corresponden a ejemplos contraintuitivos; estos problemas serán retomados en el último capítulo, y serán resueltos, otra vez, de la manera clásica; Es por ello que la lista de ejercicios que veremos aquí, es a su vez, la lista de ejercicios del último capítulo.

3.1 El Ejemplo de la Bolsa de dos Monedas

Este ejemplo tiene dos objetivos principales. Primero, mostrar que la independencia de dos Eventos Aleatorios queda determinada por las probabilidades asignadas. Segundo, mostrar que el método empírico de asignar probabilidades no siempre resulta ser correcto y adecuado.

Este Evento Aleatorio consiste en extraer una moneda de una bolsa que contiene dos de ellas, para luego, lanzarla una cierta cantidad de veces. Una de las dos monedas es equilibrada y la otra tiene el doble de probabilidad de obtener cara que sello; Nos parece que los lanzamientos son independientes; Lo cierto es que no lo son.

Analicemos el caso de dos lanzamientos en secuencia: Sacar una moneda de la bolsa, luego el primer lanzamiento y finalmente el segundo lanzamiento. Lo que resulta cierto, y podría ser una de las causas de confusión, es que los lanzamientos son *intercambiables*. Es decir, el Evento Aleatorio, formado por los dos lanzamientos en orden inverso resulta ser igual (lo mismo para cualquier par de lanzamientos). En particular, las probabilidades de obtener cara en ambos lanzamientos es la misma.

Antes de explicarlo en términos matemáticos intentemos hacerlo al estilo de Laplace. Si los Eventos fueran independientes, el resultado del primer lanzamiento no debiera cambiar la información sobre el posible resultado del segundo. Inicialmente, no hay razón alguna para suponer que la moneda extraída de la bolsa tiene más posibilidades de ser la equilibrada que de ser la cargada y, por lo tanto, es equiprobable. Si el primer lanzamiento fue cara seguir asumiendo que hay probabilidad $\dfrac{1}{2}$ de que sea la moneda equilibrada no es razonable. Lo natural es pensar que ahora es más probable que sea la moneda cargada.

De hecho la probabilidad de obtener cara con la moneda equilibrada es $\dfrac{1}{2}$ y con la moneda cargada es $\dfrac{2}{3}$. Esto nos lleva a concluir, informalmente, que la probabilidad de que la moneda sea la cargada es proporcional a esta relación. Es decir, $\dfrac{\frac{2}{3}}{\frac{2}{3}+\frac{1}{2}} = \dfrac{4}{7}$.

La probabilidad de que la moneda salga cara en el segundo lanzamiento, sabiendo que salió cara primero es el promedio ponderado de las probabilidades de obtener cara con cada moneda, es decir, $\dfrac{4}{7}\cdot\dfrac{2}{3}+\dfrac{3}{7}\cdot\dfrac{1}{2} = \dfrac{25}{42}$. Suponiendo sello en el primer lanzamiento podemos deducir que la probabilidad de que la moneda sea la cargada, es $\dfrac{2}{5}$ y por tanto la probabilidad de que el segundo lanzamiento sea sello sabiendo que salió sello el primero, es $\dfrac{2}{5}\dfrac{1}{3} + \dfrac{3}{5}\dfrac{1}{2} = \dfrac{13}{30}$.

Este tipo de razonamiento y cálculos es rápido y simple para alguien que tiene experiencia y práctica en la teoría, pero, para los que no la tienen resulta extremadamente difícil de seguir. Es, por tanto, un mal camino para un enfoque pedagógico apelar a tanto sentido común al mismo tiempo.

Veamos cómo resolverlo con nuestros discos. Si llamamos A y B a las monedas equilibrada y cargada respectivamente, por lo ya analizado ambas tienen probabilidad $\dfrac{1}{2}$ de salir de la bolsa. El modelo de probabilidades para dos lanzamientos de una moneda conocida ya lo hicimos en el Capítulo anterior. Las Tablas Marginal y Conjunta son:

A				B			
1/2				1/2			
CC	CS	SC	SS	CC	CS	SC	SS
1/4	1/4	1/4	1/4	4/9	2/9	2/9	1/9

	A	B
CC	1/8	4/18
CS	1/8	2/18
SC	1/8	2/18
SS	1/8	1/18

La Tabla Marginal que nos interesa es la de los dos lanzamientos de la moneda sin determinar cuál moneda es. Esto se obtiene sumando las dos columnas de la Tabla Conjunta (por Probabilidades Totales), para luego trasponerla, es decir:

CC	CS	SC	SS
25/72	17/72	17/72	13/72

Finalmente la Tabla Condicional y la Tabla Conjunta de estos dos Eventos Aleatorios son:

C1		S1	
21/36		15/36	
C2	S2	C2	S2
25/42	17/42	17/30	13/30

	C1	S1
C2	25/72	17/72
S2	17/72	13/72

Esto muestra que los lanzamientos no son independientes y la probabilidad de obtener cara en el primer lanzamiento es $\frac{21}{36}$. La Tabla Conjunta de Bayes es igual a la Tabla Conjunta, es decir, los Eventos son intercambiables y ambos lanzamientos tienen la misma probabilidad de salir cara o sello. Esto se produce porque la Tabla Conjunta es simétrica que no resulta lo mismo que decir que las filas son iguales, puesto que esto corresponde al caso en que son independientes.

Si aplicamos el método empírico para determinar la probabilidad de obtener cara en cualquier lanzamiento, vamos a realizar el Evento una gran cantidad de veces. Nos surge, en este caso, una duda respecto a cómo realizar el Evento muchas veces. O bien sacamos una moneda una vez y la lanzamos muchas veces, o sacamos cada vez una moneda, la lanzamos, la devolvemos y la volvemos a lanzar y así sucesivamente.

Esto es muy importante porque el segundo método sí genera Eventos Independientes, puesto que siempre se conoce la moneda pero no intercambiables, puesto que en cada lanzamiento la moneda puede ser diferente a la de otro lanzamiento. Observemos que ésta es la situación inversa a nuestro caso en que son intercambiables, pero no independientes.

Si consideramos la frecuencia relativa de caras en, digamos, 100 lanzamientos de una misma moneda, obtendremos un valor cercano a $\frac{1}{2}$, cuando la moneda sea la equilibrada y cercano a $\frac{2}{3}$, cuando la moneda sea la cargada. Si realizamos esta secuencia de 100 lanzamientos, muchas veces sacando la moneda la primera vez obtendremos una frecuencia relativa que será el promedio de esos dos casos $\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot\frac{2}{3}=\frac{7}{12}$.

Invito a los lectores a verificar lo dicho anteriormente de manera empírica y verán entonces, que la frecuencia relativa no siempre les dará la probabilidad correcta si no hay independencia aún si son intercambiables.

Nota pedagógica: Esta variante resulta ser una buena innovación pedagógica como actividad de lanzamiento de monedas. Constatar que, si se conoce la moneda, la frecuencia relativa nos da una buena aproximación de la probabilidad de obtener cara, es un clásico cuya utilidad me parece limitada y su nivel de entretención, peor aún.

3.2 Las Reinas en el Tablero de Ajedrez

El siguiente ejemplo tiene por objetivo mostrar la utilidad de modelar Eventos Aleatorios de manera Compuesta y que, a veces, los Espacios Muestrales adecuados provienen de las preguntas a responder y no del Evento Muestral natural para el Evento.

Se tiene un tablero de ajedrez y se colocan dos reinas en el tablero (en posiciones distintas) ¿cuál es la probabilidad de que se amenacen?, ¿Cuál es la probabilidad de que la primera reina esté en la Región 3 (ver figura 3.1), dado que las reinas se amenazan?

Un tablero de ajedrez es un tablero de 8 filas y 8 columnas. De las reglas del ajedrez sólo es necesario saber que la Reina puede desplazarse horizonal, vertical y en diagonal todo lo que quiera y esto corresponde a las regiones que *amenaza*.

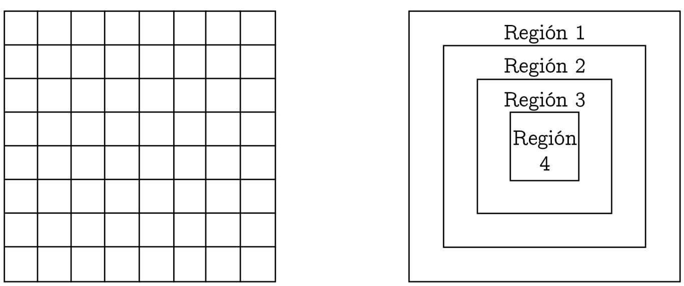


FIGURA 3.1. Regiones

El Evento Aleatorio corresponde a colocar en secuencia dos Reinas en un tablero en casillas distintas. La estructura del evento entrega naturalmente dos Eventos Aleatorios en Secuencia: la colocación de la primera y la segunda reina, con 64 posibilidades todas igualmente probables para el primero y 63 posibilidades igualmente probables para el segundo. Sin embargo, para construir la Tabla Marginal es necesario saber cuántas casillas amenaza una Reina en el tablero. Después de un poco de reflexión se aprecia que esto depende de la posición de la Reina en el tablero, de hecho esto forma 4 regiones diferentes que amenazan 21,23,25 y 27 casillas respectivamente (ver figura 3.1).

Nota: Es por esto que, en ajedrez, siempre se intenta dominar el centro del tablero.

En vista de lo anterior, en vez de considerar el primer Evento con 64 resultados posibles equiprobables, éste Primer Evento corresponde a lanzar la Reina en alguna de las 4 regiones. La probabilidad de que la Reina caiga en alguna de estas regiones es proporcional al número de casillas que tiene cada una que son 28, 20, 12 y 4, respectivamente. La información condicional sobre la probabilidad de que la segunda Reina caiga en la región amenazada por la primera, ya quedó claramente establecida (de hecho así se decidieron las regiones) por lo que la Tabla Marginal y la Condicional son:

Región 1		Región 2		Región 3		Región 4	
28/64		20/64		12/64		4/64	
SI	NO	SI	NO	SI	NO	SI	NO
21/63	42/63	23/63	40/63	25/63	38/63	27/63	36/63

	Región 1	Región 2	Región 3	Región 4
SI	$^{28 \cdot 21}/_{64 \cdot 63}$	$^{20 \cdot 23}/_{64 \cdot 63}$	$^{12 \cdot 25}/_{64 \cdot 63}$	$^{4 \cdot 27}/_{64 \cdot 63}$
NO	$^{28 \cdot 42}/_{64 \cdot 63}$	$^{20 \cdot 40}/_{64 \cdot 63}$	$^{12 \cdot 38}/_{64 \cdot 63}$	$^{4 \cdot 36}/_{64 \cdot 63}$

Reetiquetando obtenemos que la probabilidad de que se amenacen es:

$$\frac{28}{64} \cdot \frac{21}{63} + \frac{20}{64} \cdot \frac{23}{63} + \frac{12}{64} \cdot \frac{25}{63} + \frac{4}{64} \cdot \frac{27}{63} = \frac{13}{36}.$$

La segunda pregunta se resuelve con el Disco de Bayes asociado:

	SI	NO
Región 1	$^{28 \cdot 21}/_{64 \cdot 63}$	$^{28 \cdot 42}/_{64 \cdot 63}$
Región 2	$^{20 \cdot 23}/_{64 \cdot 63}$	$^{20 \cdot 40}/_{64 \cdot 63}$
Región 3	$^{12 \cdot 25}/_{64 \cdot 63}$	$^{12 \cdot 38}/_{64 \cdot 63}$
Región 4	$^{4 \cdot 27}/_{64 \cdot 63}$	$^{4 \cdot 36}/_{64 \cdot 63}$

Este mismo tipo de razonamiento se puede seguir recursivamente para n Reinas sobre el tablero. Se ve claramente que si $n > 8$ entonces la probabilidad será 1. De esta manera, se puede probar, usando las probabilidades, que existen configuraciones de 8 Reinas, en el tablero, que no se amenazan *sin* dar una solución explícita a este problema.

Esto se relaciona con el famoso *Problema de las Ocho Reinas* del que damos algunos detalles históricos. El Problema consiste en encontrar todas las diferentes formas de colocar ocho reinas sobre un tablero de ajedrez, de modo que no se amenacen entre sí.

Este problema transcendió a los círculos matemáticos de mediados del siglo XIX; el alemán Max Bezzel lo planteó en la revista Berliner Schachzeitung en 1848; el problema llegó a cautivar la atención de Gauss; el estadounidense Sam Lloyd, máxima autoridad de la *Matemática Recreativa* de la época, también lo estudió con detenimiento. Sin embargo, fue el matemático ciego Dr. Nauck el primero que encontró todas las soluciones en 1850 (ver figura 3.2).

3.2.1 Ejercicios Propuestos

1. Resuelva el mismo problema para un tablero de 2x2, 3x3, 4x4, 5x5, 6x6 y 7x7.

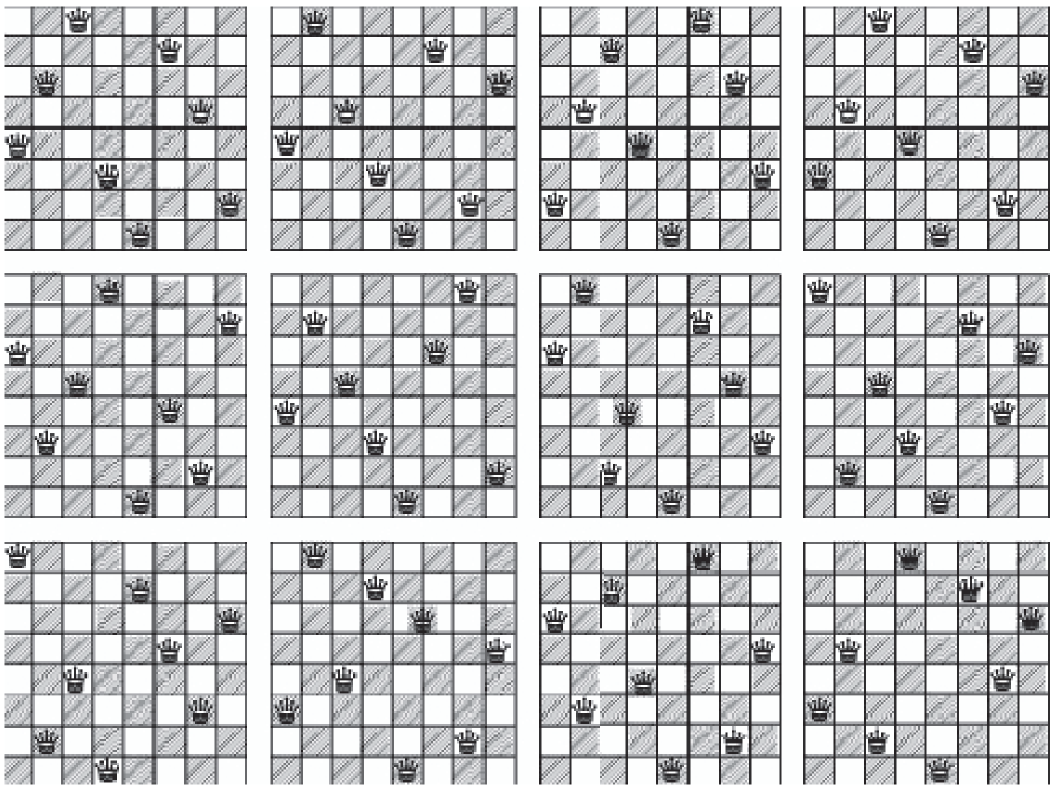


FIGURA 3.2. Total de soluciones al problema de las 8 reinas diferentes en reflexión y rotación

2. Calcule la probabilidad de que el siguiente par de piezas estén amenazadas al ser colocadas en el tablero:
   a) Dos Alfiles (se mueven por las diagonales)
   b) Dos Torres (se mueven por la horizontal o la vertical)
   c) Dos Reyes (en todas direcciones pero 1 solo paso).
3. Sabiendo que las dos Reinas están amenazadas, calcule la probabilidad de que se encuentren en la misma región para cada una de las regiones 1,2,3 ó 4.

4. Teniendo en cuenta el contexto del problema anterior, calcule la probabilidad de que se encuentren en una misma región (sin saber cuál).

3.3 El Problema de la Mesa Redonda

El objetivo del siguiente ejemplo es entregar un caso que se puede resolver, tanto con probabilidades, como por Combinatoria y cuya dificultad principal proviene de seleccionar los Eventos Compuestos adecuados.

Se tiene una mesa redonda con 16 sillas en las que se sientan tres personas ¿cuál es la probabilidad de que no haya dos de ellos contiguos?.

Es claro que este Evento Compuesto se puede modelar con 3 Eventos en Secuencia que consisten en tres personas sentándose. Como ya es habitual en nuestro caso, lo vamos a modelar, sólo con dos, combinando la ubicación de las dos primeras personas.

Al sentar a las primeras dos personas hay tres situaciones distintas a considerar para resolver nuestro problema. Si están contiguos, a 1 silla de distancia o a más de 1 silla de distancia. En efecto, para que el tercero quede contiguo tiene 3 posibilidades si las dos primeras personas están a 1 silla de distancia y 4 posibilidades si están a más de 1 silla de distancia (ver las X's en la figura 3.3).

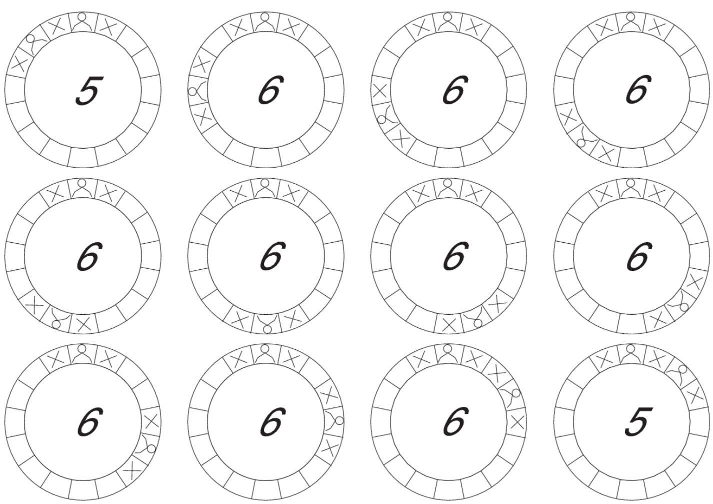


Figura 3.3. Casos diferentes de dos personas no contiguas sentadas

Entonces, el primer Evento Aleatorio se construye con estos 3 resultados y no con las $16 \cdot 15$ posibilidades de sentarlos a los dos. Es decir, quedan las dos personas contiguas, no contiguas a una silla de distancia y finalmente no contiguas a más de una silla de distancia. Las probabilidades de cada uno de estos resultados son simples. Con la primera persona sentada de los 15 puestos restantes hay 2 puestos para quedar contiguo, 2 para quedar a 1 silla de distancia y 11 para quedar a más de 1 silla de distancia. Para acotar notación notamos 2C si quedan dos contiguos, NC1 si quedan no Contiguos a 1 silla y NC2 si quedan no Contiguos a más de 1 silla. La Tabla Marginal será:

2C	NC1	NC2
2/15	2/15	11/15

El segundo Evento Aleatorio consiste en sentar a la tercera persona y determinar si alguno de los 3 quedó o no contiguo, que notaremos como C o NC, respectivamente. Por lo ya discutido, es directo obtener el Disco Condicional para cada uno de los casos, por lo que la Tabla Marginal y la Tabla Conjunta nos dan:

2C		NC1		NC2	
2/15		2/15		11/15	
C	NC	C	NC	C	NC
1	0	3/14	11/14	4/14	10/14

	2C	NC1	NC2
C	2/15	$(2/15)\cdot(3/14)$	$(11/15)\cdot(4/14)$
NC	0	$(2/15)\cdot(11/14)$	$(11/15)\cdot(10/14)$

El caso que nos interesa es que ninguno de los 3 esté contiguo y reetiquetando la respuesta a nuestra pregunta es $\frac{2}{15}\frac{11}{14}+\frac{11}{15}\frac{10}{14}=\frac{22}{35}$.

Este ejemplo será resuelto en un contexto más general en el capítulo de Combinatoria, donde se darán algunas variantes como ejercicios propuestos.

3.4 El Problema de los Hijos

El siguiente Evento es un ejemplo de cómo desarrollar la intuición con la Regla de Laplace sin un razonamiento lógico, nos lleva a conclusiones erradas. Este caso, en particular, es uno de los ejemplos más utilizados en la literatura por su simpleza de planteamiento.

Un hombre visita a un matrimonio que tiene dos hijos. Uno de los hijos, un niño (varón), entra en la sala, ¿cuál la probabilidad de que el otro sea también varón?, si:

1. se sabe que el otro hijo (o hija) es menor;
2. no se sabe nada del otro hijo.

Sin mayor información se asume que la probabilidad de que un hijo cualquiera sea niño o niña es inicialmente $\frac{1}{2}$. Por lo tanto parece natural asumir que la probabilidad pedida es $\frac{1}{2}$. Lo cierto es que esto no resulta ser correcto.

Tenemos dos Eventos Aleatorios iguales que corresponden al nacimiento de cada uno de los hijos. Al considerarlo como disco obtenemos 4 regiones iguales etiquetadas como hh, hm, mh y mm. Donde hm representa "primero nació un hijo y luego una hija". En el caso (1), como en la sala vemos a un varón y sabemos que el otro hijo(a) es menor, estamos eliminando los resultados hm y mm que corresponden a los casos en que el segundo hijo es mujer. Esta simplificación deja las probabilidades de cada uno de los dos resultados restantes en $\frac{1}{2}$. Reetiquetando los dos resultados, al borrar la segunda palabra de cada etiqueta, nos damos cuenta que una representa el segundo hijo hombre y mujer. Esto corresponde a nuestra intuición y es consistente porque hemos fijado el orden de los hijos.

Si estamos en la situación (2) sólo sabemos que uno de los dos hijos es un varón que corresponde a eliminar únicamente la región mm. El nuevo disco tendrá tres regiones hh, hm y mh con igual área, es decir $\frac{1}{3}$. De estos tres resultados, dos representan el caso de que el otro hijo sea del sexo opuesto y sólo 1 a que el otro hijo sea del mismo sexo y, por tanto, la probabilidad es $\frac{1}{3}$.

3.5 El Juego de las Tarjetas

En este caso presentamos un ejemplo que muestra lo importante que es la representación de un Evento Aleatorio al momento de plantearlo para desarrollar la intuición.

Se tienen tres tarjetas, una roja por ambos lados, una blanca por ambos lados y una blanca por un lado y roja por el otro. Se toma una tarjeta al azar y se coloca sobre la mesa. Sabiendo que la cara superior es roja, ¿cuál es la probabilidad que la otra también lo sea?.

Habitualmente, al plantear este problema la respuesta es $\frac{1}{2}$, incluso entre alumnos de prestigiosas universidades que ya han tenido un curso básico de probabilidades. Curiosamente un replanteamiento simple produce una respuesta diferente.

Por ejemplo, si se tienen tres cajones cerrados, uno de ellos con dos monedas de oro, otro con dos monedas de plata y el último con una moneda de plata y otra de oro y se saca una moneda al azar de uno de los tres cajones y ésta es de plata ¿cuál es la probabilidad que la moneda restante lo sea también?

Es poco habitual recibir la misma respuesta que en el caso anterior, aún cuando es claro que son el mismo problema reemplazando cajones por tarjetas y lados por monedas.

El primer evento está asociado al experimento de seleccionar la carta, cuya tabla es bastante simple y que etiquetamos como $C1$, $C2$ y $C3$ para las cartas roja por ambos lados, blanca por ambos lados y roja por un lado, blanca por el otro. El segundo evento está asociado a colocar la carta sobre la mesa y consideramos 4 resultados posibles RR, RB, BR y BB en que la primera etiqueta corresponde a la cara expuesta y la segunda a la cara oculta de la carta. Las Tablas Condicional y Conjunta dan:

$C1$				$C2$				$C3$			
1/3				1/3				1/3			
RR	RB	BR	BB	RR	RB	BR	BB	RR	RB	BR	BB
1	0	0	0	0	0	0	1	0	1/2	1/2	0

	$C1$	$C2$	$C3$
RR	1/3	0	0
RB	0	0	1/6
BR	0	0	1/6
BB	0	1/3	0

Las Tablas de Bayes asociadas son:

	RR	RB	BR	BB
$C1$	1/3	0	0	0
$C2$	0	0	0	1/3
$C3$	0	1/6	1/6	0

RR			RB			BR			BB		
1/3			1/6			1/6			1/3		
$C1$	$C2$	$C3$	$C1$	$C2$	$C3$	$C1$	$C2$	$C3$	$C1$	$C2$	$C3$
1	0	0	0	0	1	0	0	1	0	1	0

Nota: Es importante advertir que al invertir la causalidad, una vez que se fija el resultado de la carta sobre la mesa, ¡la carta seleccionada queda fija también y no hay realmente azar en el segundo evento!.

Si nos interesa sólo el resultado de las cartas, consideramos sobre la primera fila de la Tabla Marginal de Bayes y simplificamos respecto a obtener rojo en la cara superior, obteniendo:

RR	RB
2/3	1/3

Entonces, la probabilidad de que la otra cara sea roja es el doble a que sea blanca.

3.5.1 Ejercicios Propuestos

1. ¿Es posible obtener igual probabilidad de obtener una cara inferior roja o blanca si la cara superior es roja, cambiando las probabilidades de obtener cada una de las cartas? Si es así, encuentre esas probabilidades.

2. En el ejercicio anterior, si se logra igual probabilidad cuando sale roja en la cara superior, ¿esto también asegura lo mismo para el caso en que la cara superior es blanca?

3. ¿Es posible obtener igual probabilidad de obtener una cara inferior roja o blanca si la cara superior es roja cambiando las probabilidades de que las cartas caigan de un lado o del otro? Si es así, encuentre esas probabilidades.

3.6 El problema del Falso Positivo

Este último problema pretende mostrar una aplicación más realista y seria que los problemas precedentes. La idea es aclarar que incluso a este nivel podemos plantear y resolver problemas contingentes y de interés general.

Una joven estudiante se toma un test de embarazo y le da positivo. Para estar segura se toma otro test igual y esta vez le da negativo. ¿Cuál es la probabilidad de que esté embarazada?, ¿cuál sería la probabilidad de estar embarazada si el test le diera positivo dos veces seguidas?, ¿qué pasa si el test le diera negativo una vez y luego positivo?

Al revisar la caja, la estudiante ve que el test indica que da positivo falsamente con probabilidad menor o igual a un 0.1 % y da negativo falsamente con probabilidad menor o igual 0.01 %.

Nota: En el área de medicina, en estos casos, se considera peor dar un diagnóstico negativo por error, que dar un diagnóstico positivo por error, lo que es muy pero muy razonable.

A pesar de lo simple que resulta enunciar este problema y lo relevante que puede resultar como aplicación en nuestro mundo de hoy, se genera mucha confusión al intentar resolverlo.

En esta situación tenemos dos Eventos Aleatorios, primero si la alumna está o no embarazada con una probabilidad que no conocemos y que notaremos p, segundo, si el test se equivocó o no. La confusión proviene de confundir entre que el test diga positivo y que la alumna esté embarazada.

Notaremos como P(Positivo) y N(Negativo) a los resultados del test y SI y NO si la estudiante está o no embarazada.

La Tabla Condicional y Conjunta son:

SI		NO	
p		$1-p$	
P	N	P	N
0.9999	0.0001	0.001	0.999

	SI	NO
P	$0{,}9999 \cdot p$	$0{,}001 \cdot (1-p)$
N	$0{,}0001 \cdot p$	$0{,}999 \cdot (1-p)$

La Tabla de Bayes asociada es:

	P	N
SI	$0{,}9999 \cdot p$	$0{,}0001 \cdot p$
NO	$0{,}001 \cdot (1-p)$	$0{,}999 \cdot (1-p)$

P		N	
$0{,}001 + 0{,}9989 \cdot p$		$0{,}999 - 0{,}9989 \cdot p$	
SI	NO	SI	NO
$0{,}9999 \cdot p / 0{,}001 + 0{,}9989 \cdot p$	$0{,}001 \cdot (1-p) / 0{,}001 + 0{,}9989 \cdot p$	$0{,}0001 \cdot p / 0{,}999 - 0{,}9989 \cdot p$	$0{,}999 \cdot (1-p) / 0{,}999 - 0{,}9989 \cdot p$

Entonces la probabilidad de que la estudiante esté embarazada, dado que el test dio positivo, es $\dfrac{0{,}9999 \cdot p}{0{,}001 + 0{,}9989 \cdot p}$ y dado que el test dio negativo es, $\dfrac{0{,}0001 \cdot p}{0{,}999 - 0{,}9989 \cdot p}$.

Siempre resulta interesante interpretar los resultados en los *Casos Límite* para ver si los cálculos tienen sentido. En este contexto estamos hablando de los casos en que el azar es mínimo que es $p = 1$ y $p = 0$.

Cuando $p = 1$, la probabilidad de que la estudiante esté embarazada dado que el test dio positivo, es 1 al igual que si da negativo. Esto es claro: no se pueden hacer milagros, los resultados del test no pueden cambiar esta situación. Ahora bien, la probabilidad de que el test resulte positivo es igual a 0.9999 que corresponde a la probabilidad de que el test correctamente dé positivo y la probabilidad de que el test dé negativo es igual a 0.0001 que corresponde a la probabilidad de que el test dé incorrectamente positivo.

Cuando $p = 0$ la probabilidad de que la estudiante esté embarazada, dado que el test dio positivo, es 0 y lo mismo si el test da negativo. Una vez más no hay milagros. La probabilidad de que el test dé positivo es 0.001 que corresponde a la probabilidad de que el test dé positivo erróneamente y la probabilidad de que el test dé negativo es 0.999 que que corresponde a la probabilidad que el test dé negativo correctamente.

Entonces, la estudiante tiene que definir según su información cual es su estimación de la probabilidad de estar embarazada que no será ni 0 ni 1 puesto que no tomaría el test si no tuviera dudas o si supiera que está embarazada.

Se considera que la probabilidad de que una joven quede embarazada si tiene sexo sin protección (SP) es de alrededor de un 85 %, con Espermicida (E) alrededor de un 18 %, Diafragma (D) 16 % y con condón (C) un 2 %.

Si tabulamos las probabilidades de que la estudiante esté embarazada, si el test da positivo o negativo para estos casos (aproximando al sexto decimal) nos da:

	Si dado P	Si dado N
SP	0,999824	0,000567
E	0,9954647	0,000022
D	0,9947769	0,000019
C	0,9532844	0,000002

Si la estudiante toma el test dos veces ahora hay 4 resultados posibles que notaremos PP,PN,NP y PP. Intuitivamente los dos tests no pueden ser independientes entre sí. Si el primero da positivo es más probable que el segundo también lo haga. Pero dos test si son independientes al tomarlos con una persona embarazada o una que no lo está (es decir sabiendo su estado de gravidez), puesto que funcionan de la misma manera ante esa muestra.

La Tabla Condicional y Conjunta son:

SI				NO			
p				$1-p$			
PP	PN	NP	NN	PP	PN	NP	NN
$0{,}9999^2$	$0{,}00009999$	$0{,}00009999$	$0{,}0001^2$	$0{,}001^2$	$0{,}000999$	$0{,}000999$	$0{,}999^2$

	SI	NO
PP	$0{,}9999^2 \cdot p$	$0{,}001^2 \cdot (1-p)$
PN	$0{,}00009999 \cdot p$	$0{,}000999 \cdot (1-p)$
NP	$0{,}00009999 \cdot p$	$0{,}000999 \cdot (1-p)$
NN	$0{,}0001^2 \cdot p$	$0{,}999^2 \cdot (1-p)$

La Tabla Conjunta de Bayes:

	PP	PN	NP	NN
SI	$0{,}9999^2 \cdot p$	$0{,}00009999 \cdot p$	$0{,}00009999 \cdot p$	$0{,}0001^2 \cdot p$
NO	$0{,}001^2 \cdot (1-p)$	$0{,}000999 \cdot (1-p)$	$0{,}000999 \cdot (1-p)$	$0{,}999^2 \cdot (1-p)$

No es necesario escribir completa la Tabla de Bayes Marginal una vez que se sabe cómo se realizan los cálculos para una pregunta en particular. Por ejemplo, la probabilidad de que la estudiante esté embarazada, dado que el test dio positivo dos veces, será:

$$\frac{0{,}9999^2 \cdot p}{0{,}001^2 + 0{,}99979901 \cdot p}$$

Para el caso en que el test dio positivo y luego negativo será:

$$\frac{0{,}00009999 \cdot p}{0{,}000999 - 0{,}00089901p.}$$

Si tabulamos como antes para estos dos casos tenemos:

	Si dado PP	Si dado PN
SP	0,999999823	0,361910054
E	0,999995444	0,021498649
D	0,999994749	0,018708113
C	0,999950993	0,002038491

3.7 Ejercicios

1. Se saca dos veces una bola de una urna con N bolas numeradas de 1 a N (siempre se devuelven las bolas). Calcular la probabilidad de que el número de la primera bola sea mayor (estricto) que el de la segunda. En general, si se sacan n bolas, calcular la probabilidad de que estén ordenadas de mayor a menor, ¿es igual a la probabilidad de que estén en algún otro orden?.

2. Al tirar un dado equilibrado, con iguales probabilidad se obtienen los números 1,2,3,4,5 ó 6. En caso de tirar dos dados la suma de los puntos obtenidos está comprendida entre 2 y 12. Tanto el 9 como el 10, a partir de los números 1,2,3,4,5,6 se puede obtener de dos formas distintas: 9=3+6=4+5, y 10=4+6=5+5. En el problema con tres dados, tanto el 9 como el 10 se obtienen de seis formas. ¿Por qué entonces, el 9 se obtiene con mayor frecuencia al tirar dos dados, y el 10 con mayor frecuencia al tirar tres?

   Nota: Este problema aparece por primera vez en el libro de G. Cardano.

3. Todos los días a las 5:00 pm Charles toma el té con su madre. Para empezar la conversación dice "creo que llueve" o bien "creo que no llueve". Charles se equivoca 1 vez de 3 cuando afirma que llueve y 1 vez de 2 cuando afirma que no llueve. Sabiendo que llueve 9 de cada 10 días y que Charles dijo que llovía calcule la probabilidad de que llueva.

4. A apuesta a B, que de un mazo de 40 cartas, entre las cuales hay 10 de cada color, extraerá 4, de forma de obtener una de cada color, ¿cuáles son las probabilidades de A sobre las de B?
   Respuesta: 1000 contra 8139.

   Nota: Este problema fue propuesto por Fermat en carta a Huygens en junio de 1656, la respuesta está en la carta a Carcavi del 6 de julio de 1656.

5. Se tienen tres discos A, B y C (ver figura 3.4).

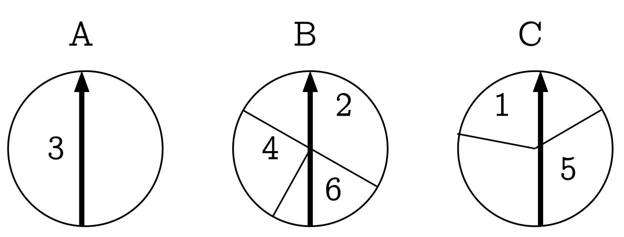


FIGURA 3.4. Discos

Diremos que un disco es mejor que otro si tiene probabilidad mayor que $\frac{1}{2}$ de obtener un valor mayor que otro. Notaremos como A, B y C a los valores aleatorios obtenidos al hacer girar las flechas de los discos A, B y C, respectivamente.

a) Muestre que A es mejor que B, B es mejor que C pero que C es mejor que A cuando las Tablas de los Discos B y C son:

2	4	6
2/3	1/6	1/6

1	5
4/9	5/9

b) ¿ Si la tabla de C es

1	5
1/3	2/3

cómo cambia lo anterior?

6. Un fugitivo debe cruzar la frontera entre dos países. Para ello dispone de 3 caminos. Los dos primeros caminos tienen un puesto de control y el tercero tiene dos puestos de control en serie, a una distancia prudente. Se asume además que cada puesto de control es independiente de los otros. Se sabe que la probabilidad de burlar un puesto de control es $p \in (0,1)$ (para cada puesto es la misma probabilidad) y que el fugitivo toma al azar uno de los tres caminos.

   a) Calcule la probabilidad de que el fugitivo cruce la frontera sin ser visto.

   b) Calcule la probabilidad de que el fugitivo haya utilizado el tercer camino, dado que pasó la frontera sin ser visto.

   c) Si los guardias de frontera deciden al azar e independientemente de la decisión del fugitivo controlar completamente uno de los tres caminos, ¿cual es la probabilidad de que el fugitivo alcance la frontera?

7. Dos personas A y B participan en un juego aleatorio, para el cual cuentan con dos dados equilibrados y una urna con n bolitas negras y b bolitas blancas. El juego comienza al lanzar ambos dados para luego sacar dos bolitas de la urna. Si la suma de los dados es mayor o igual a 7 o menor o igual que 3 entonces gana B, si se obtiene alguna bola blanca y gana A en otro caso. Si la suma de los dados es mayor que 3 y menor que 7 entonces A gana cuando se obtienen bolitas de distinto color y B gana en otro caso.

   a) Determine una relación entre n y b de modo que el juego sea justo para ambos participantes.

   b) Suponga que la urna tiene 2 bolitas blancas y 6 negras. Sabiendo que A ganó el juego, calcule la probabilidad de que la suma de los dados haya sido mayor o igual que 7.

Capítulo 4: Combinatoria Básica

La Combinatoria es el nombre que recibe el estudio de las combinaciones o agrupaciones posibles de un determinado número de elementos. En este Capítulo introduciremos las técnicas básicas de la combinatoria que no es más que una ínfima parte de esta Teoría. En estricto rigor, la combinatoria es mucho más que técnicas de conteo y es por eso que preferimos llamar a este capítulo Combinatoria Básica y no Combinatoria simplemente.

Las técnicas de conteo son anteriores incluso a la historia de la matemática. Sin siquiera saber aritmética los pastores de antaño utilizaban técnicas combinatorias para saber si perdían alguna oveja al volver a casa. El pastor simplemente llenaba su morral con piedras, una por cada oveja que salía y luego, al regresar a casa, las sacaba a medida que las ovejas regresaban. Si le sobraban piedras sabía que había perdido alguna, si no le faltaban, entonces volvió con lo que se fue y si las piedras no le alcanzaban, se había apropiado de las ovejas de alguien más.

Uno de los ejemplos más antiguos de combinatoria aparece en la obra china del I - Ching ("Libro de las transformaciones") con sus combinaciones de trigramas místicos que, según historiadores, estaba inspirado en un libro sobre variaciones escrito en el Japón del siglo XII de nuestra era. Lo cierto es que los griegos no dedicaron mucho tiempo a la combinatoria, con la honrosa excepción del estudio de los números poligonales de los pitagóricos. El Imperio Romano no tuvo globalmente gran interés en el desarrollo de las matemáticas y como curiosidad se encuentra en Boecio (siglo V) la regla para encontrar las combinaciones de n objetos tomados dos a dos. En India, en el siglo XII, ya se conocían los coeficientes binomiales y Bhaskara, un siglo antes, encontró reglas para calcular variaciones (con y sin repetición) y combinaciones con algunas aplicaciones. En el mundo medieval, por las aplicaciones a la Cábala Judía y a la Astronomía, se prestó bastante atención a la combinatoria. Pero fue Leví Ben Gerson en el siglo XIV el que realizó un estudio detallado de lo que se conoce hoy en día como permutaciones y combinaciones de un conjunto de objetos. Esto fue recuperado por Pascal y Fermat en sus estudios de los juegos de azar. Pascal fue el primero en relacionar los coeficientes binomiales con el Teorema del Binomio. Aunque Nicolo Fontano de Bresia, más conocido como Tartaglia, en su obra póstuma "Tratado general sobre el número y la medida" estudió el rectángulo aritmético que corresponde a lo que actualmente llamamos *Triángulo de Pascal* (ver figura 4.1).

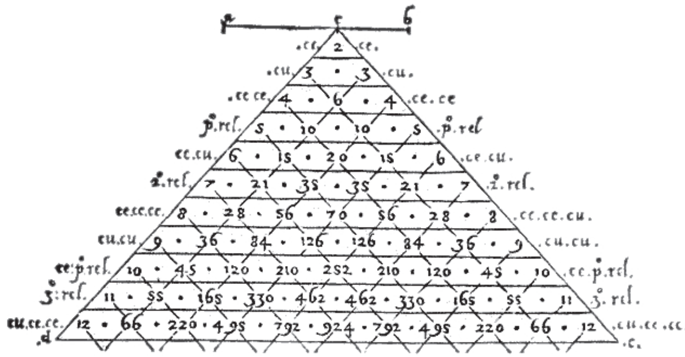


FIGURA 4.1. Rectángulo Aritmético de Tartaglia

4.1 Preliminares

El concepto más básico de la Combinatoria es la definición de *cardinal* de una agrupación de elementos que es una medida de cantidad pero no de orden. En castellano, la definición es el adjetivo o pronombre numeral que indica únicamente cantidad o número. En el caso de una cantidad finita de elementos, el cardinal corresponde al número de elementos.

La definición formal que permite definir el cardinal, para cualquier agrupación de elementos más allá de los conjuntos finitos, proviene de la idea de nuestro amigo el pastor y sus ovejas. El procedimiento del pastor para obtener el cardinal de sus ovejas era, en el fondo, establecer una función entre sus ovejas y las piedras en su bolsa.

Si seguimos el razonamiento de nuestro ancestral pastor y notamos cómo $|A|$ el cardinal del conjunto A entonces dado B otro conjunto y f una función $f : A \to B$ tendremos que:

- Si f es *sobreyectiva* entonces $|A| \geq |B|$.
- Si f es *inyectiva* entonces $|A| \leq |B|$.
- Si f es *biyectiva* entonces $|A| = |B|$.

Una definición básica de cardinal, que es suficiente para este Capítulo, es la siguiente:

- $|\phi| = 0$
- A tiene *Cardinal* N si existe una función biyectiva entre A y el conjunto $\{1, \ldots, N\}$.
- A tiene *Cardinal Numerable* si existe una función biyectiva entre A y el conjunto $\mathbb{N}$.
- A tiene *Cardinal no Numerable* si toda función sobreyectiva entre A y $\mathbb{N}$ no es biyectiva.

4.2 Principios Básicos de Conteo

Contar es tal vez la primera y la más simple de las operaciones combinatorias. En esta sección vamos a considerar cardinales finitos, y por tanto, nuestros objetos van a ser bolsas con una cantidad finita de elementos distinguibles.

Los dos principios básicos de conteo son el *Principio de Multiplicación* y el *Principio de Suma*. Informalmente podemos enunciarlos como:

Principio de Multiplicación:
Si una bolsa contiene n_1 elementos y otra bolsa contiene n_2 elementos distintos, entonces el total de parejas que se pueden formar con un elemento de la primera bolsa y otro de la segunda es $n_1 \cdot n_2$. Esto se extiende a más de dos operaciones (digamos k) de manera natural.

Nota: En lenguaje de conjuntos. Dados dos conjuntos finitos A y B, $|A \times B| = |A| \cdot |B|$.

Si volvemos a nuestra representación de árbol y consideramos k bolsas con n_1,...,n_k elementos respectivamente y construimos un camino por cada elemento de una bolsa y los combinamos en serie, entonces tenemos que el total de caminos posibles es $n_1 \cdot ... \cdot n_k$ (ver figura 4.2).

$$\left\{ \begin{array}{c} 1 \\ \vdots \\ n_1 \end{array} \right\} \rightarrow \left\{ \begin{array}{c} 1 \\ \vdots \\ n_2 \end{array} \right\} \rightarrow \dots \rightarrow \left\{ \begin{array}{c} 1 \\ \vdots \\ n_k \end{array} \right\}$$

FIGURA 4.2. Principio de Multiplicación

Dentro de las múltiples aplicaciones de este Principio, vamos a resaltar algunas que nos parece son particularmente importantes en nuestro contexto.

Mencionamos inicialmente que la cardinalidad es una medida de cantidad y no de orden, pero lo cierto es que se puede aplicar para contar la cantidad de maneras de ordenar!.

Aplicación 1. *El total de órdenes distintos de N elementos distinguibles en serie es:*

$$N! := N \cdot (N-1) \cdot \dots \cdot 2 \cdot 1.$$

En efecto un orden de los elementos en serie viene dado por asignar una posición a cada elemento. Hay N posiciones posibles para el primer elemento a ordenar, $(N-1)$ posiciones para el segundo elemento, ..., y finalmente 1 sóla posición restante para el último elemento a ordenar. Aplicando el Principio de Multiplicación se concluye.

Nota: Por conveniencia se define $0! := 1$.

Este número $N!$ se le llama el *factorial de* N y es una cantidad de gran importancia en la Combinatoria y aparecerá en la gran mayoría de las fórmulas de este capítulo.

Si tenemos una bolsa con N elementos y tenemos la opción de extraer tantos elementos como queramos estamos formando otra bolsa diferente. Habitualmente se tiende a olvidar un caso un poco especial que consiste en la bolsa vacía que tiene 0 elementos.

Aplicación 2. *El total de bolsas diferentes que se pueden formar de una bolsa que inicialmente tiene* N *elementos y de los cuales se extraen tantos como se quiera es* 2^N

En el fondo cada posible bolsa queda determinada por los elementos que fueron retirados. Entonces a cada elemento le podemos asignar dos valores, uno si fue extraído y otro si no. Una vez más, aplicando el Principio de Multiplicación, concluimos que el total de bolsas posibles de formar es 2^N.

Nota: En el lenguaje de conjuntos esto es equivalente a decir que el cardinal del conjunto de todos los subconjuntos de un conjunto de N elementos (el Conjunto Potencia) es 2^N.

El segundo principio en nuestro contexto de bolsas es bastante simple de enunciar.

Principio de Suma:
Si una bolsa tiene n_1 elementos distinguibles y otra bolsa tiene n_2 elementos distinguibles, de modo que además todos los elementos de la primera son distinguibles de los elementos de la segunda bolsa, al mezclar las dos bolsas en 1 sola obtenemos $n_1 + n_2$ elementos distinguibles.

Nota: En lenguaje de conjuntos. Dados dos conjuntos finitos y disjuntos A y B se tiene que $|A \cup B| = |A| + |B|$.

Aplicación 3. *Si lanzo un dado de seis caras o una moneda, entonces los resultados posibles son* $6 + 2 = 8$.

4.3 El Problema de la Mesa Redonda II

En el Capítulo anterior planteamos el problema de sentar 3 personas en una mesa redonda con 16 sillas (sección 3.3). Vamos a retomar este problema pero para cualquier cantidad de sillas mayor que 6.

¿De cuántas maneras distintas se pueden sentar 3 personas en una mesa redonda con $N \geq 6$ sillas?. De este total, ¿cuántas de ellas hacen que no haya 2 de ellos contiguos?

El total de maneras posibles se calcula directamente por el Principio de Multiplicación. Hay N lugares para sentar a la primera persona, pero solo $N-1$ lugares para sentar a la segunda persona y finalmente $N-2$ lugares para sentar a la tercera y última persona. Esto da un total de $N \cdot (N-1) \cdot (N-2)$.

Para lograr que no haya dos contiguos, hay N lugares para sentar a la primera persona y de los $N-1$ lugares restantes, hay solo $N-3$ lugares para sentar a la segunda porque los asientos contiguos al asiento de esta primera persona no pueden ser ocupados.

Ahora bien, supongamos que ubicamos a la segunda persona a una silla de distancia de la primera persona. Claramente esto lo podemos hacer de 2 formas diferentes (a la derecha o a la izquierda de ésta). En consecuencia, hay 5 posiciones para sentar a la tercera persona que la dejan contigua a alguna de las otras dos y, por tanto, $N-5$ asientos para sentar a la tercera persona sin que quede contigua a las otras dos.

Si ubicamos a la segunda persona a más de una silla de distancia de la primera, habrán 6 asientos que dejan a la tercera persona contigua a alguna de las otras dos.

En resumen, luego de ubicar a la primera persona, hay 2 lugares para los cuales hay $N-5$ asientos para sentar a la tercera persona sin que queden contiguos y $N-5$ asientos para los cuales hay $N-6$ maneras de sentar a la tercera persona, de manera que no queden contiguos (ver figura 4.3).

$$\begin{array}{cccc}
 & & 2 & \longrightarrow \; N-5 \\
 & \nearrow & & \\
N & & & \\
 & \searrow & & \\
 & & N-5 & \longrightarrow \; N-6
\end{array}$$

Figura 4.3. Árbol de casos no contiguos

En consecuencia, por los Principios de Suma y Multiplicación, hay $N \cdot 2 \cdot (N-5) + N \cdot (N-5) \cdot (N-6) = N \cdot (N-4) \cdot (N-5)$ maneras distintas de lograr que no queden dos contiguos.

Al considerar todas las configuraciones posibles como igualmente probables, obtendremos que la probabilidad de que no queden contiguos es:

$$\frac{(N-4) \cdot (N-5)}{(N-1) \cdot (N-2)}.$$

En la figura 3.3 de la sección 3.3 se grafican casos de dos personas sentadas no contiguas en el caso $N = 16$. El número al centro de los discos de la figura muestra cuántas sillas no son posibles para la tercera persona.

Nota: Para el caso $N = 16$ la probabilidad de que queden contiguos da $\dfrac{12 \cdot 11}{15 \cdot 14} = \dfrac{22}{35}$ que coincide con el resultado obtenido en la sección 3.3.

4.3.1 Ejercicios Propuestos

Resuelva el mismo problema para las siguientes distribuciones de sillas:

1. Las sillas en fila india.

2. $2N$ sillas en dos filas indias separadas de N sillas.

3. $2N - 1$ sillas en cruz con N filas en horizontal y vertical.

4. $2N$ sillas que están dispuestas en 2 círculos separados.

5. $2N - 1$ sillas que forman el símbolo ∞.

4.4 Ejercicios de los Principios de Suma y Multiplicación

En esta sección damos un pequeño listado de ejercicios a resolver aplicando los Principios de Suma y Multiplicación. A modo de ejemplo, damos la resolución detallada de los primeros.

1. Hallar el número de palabras, con o sin sentido, de tres letras distintas que pueden formarse con las letras a, b, c, d, e y f.

   Respuesta: Para la primera tenemos 6 posibilidades; para la segunda letra tenemos 5 para la última 4. Por el Principio de Multiplicación hay $6 \cdot 5 \cdot 4 = 120$ palabras posibles.
2. Un curso está formado por 12 niños y 4 niñas. Si se escogen tres estudiantes del curso al azar, ¿cuál es la probabilidad de que sean todos niños?

   Respuesta: Por el Principio de Multiplicación el total de maneras de seleccionar a 3 estudiantes al azar es $16 \cdot 15 \cdot 14$. El total de maneras de seleccionar al primer estudiante como niño son 12. Si el primero seleccionado es un niño, el total de posibilidades de seleccionar a otro niño es 11; finalmente, si los primeros dos escogidos son niños, las posibilidades en que el tercero sea niño es 10. Una vez más por el Principio de Multiplicación el total de posibilidades de obtener 3 niños es $12 \cdot 11 \cdot 10$. Entonces, al considerar favorables sobre posibles como la probabilidad obtenemos que la probabilidad de que sean todos niños es:

$$\frac{16 \cdot 15 \cdot 14}{12 \cdot 11 \cdot 10} = \frac{11}{28},$$

3. A un jugador le reparten 5 cartas de una baraja corriente de 52 cartas. Calcule la probabilidad de obtener:
   a) *Par*=Exactamente dos cartas iguales.
   b) *Doble Par*=Exactamente dos pares.
   c) *Trio*=Exactamente tres cartas iguales.
   d) *Color*=Todas de la misma pinta.
   e) *Poker*=Cuatro cartas iguales.
   f) *Escalera*=5 cartas en orden ascendente (**Nota:** Esto es cíclico, es decir, vale $K - A - 2 - 3 - 4$).
   g) *Escalera Real*=Escalera todas de la misma pinta.
   h) *Full*=Un trío y un par.

   Respuesta: Veamos cómo resolver el caso de Color:

   Por el Principio de Multiplicación el total de maneras de repartir 5 cartas es $52 \cdot 51 \cdot 50 \cdot 49 \cdot 48$. Como hay 13 cartas de cada pinta, el total de maneras de que las 5 sean de la misma pinta es $13 \cdot 12 \cdot 11 \cdot 10 \cdot 9$. La probabilidad pedida será entonces:

$$\frac{13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48} = \frac{33}{6640}.$$

4. ¿Cuál es la probabilidad de obtener 3 caras en 5 lanzamientos de una moneda equilibrada?
5. Si se tira un dado de 6 caras 5 veces, ¿cuál es el la cantidad de apariciones del seis más probable?
6. Se saca una letra al azar de “assisine” y una letra al azar de “assassin”, ¿cuál es la probabilidad de que las letras extraídas coincidan?
7. Se toman dos números enteros al azar entre 0 y 10, ¿cuál es la probabilidad de que la suma de los números sea par?, ¿cómo cambia la respuesta si los números están entre 0 y 9?

4.5 Urnas y Bolitas

En esta sección se introducen las principales técnicas combinatorias de conteo. Siguiendo con el enfoque metodológico utilizado hasta ahora, vamos a considerar como representación principal lanzar una cantidad finita de bolitas, distinguibles o no, al interior de una cantidad finita de urnas (o casillas) distinguibles (numeradas por ejemplo). El experimento consiste en lanzar *de todas las maneras posibles* las bolitas, una a una, en las casillas. Sin embargo, lo que se contará no será este total, sino la *cantidad de configuraciones distinguibles de urnas llenas con bolitas.* (ver figura 4.4).

Es importante definir con precisión lo que se entiende por configuraciones distinguibles de urnas llenas con bolitas. Dos configuraciones son distinguibles si se puede,

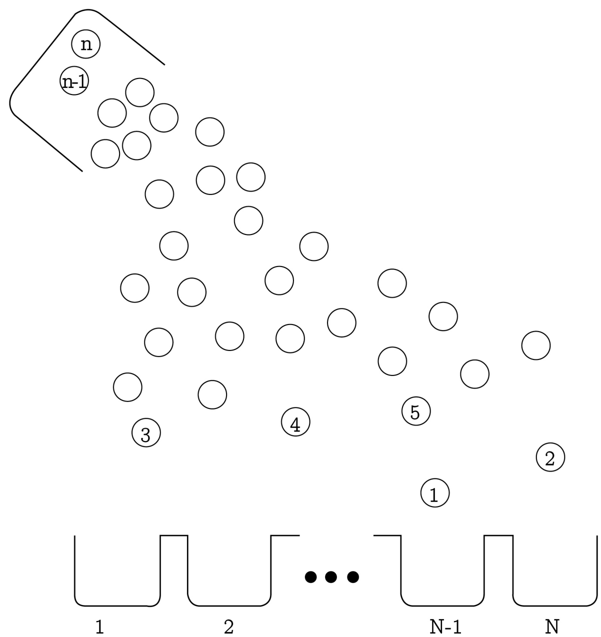


Figura 4.4. Urnas y Bolitas

entre las dos, encontrar alguna diferencia en las bolitas contenidas en alguna urna. Dependiendo de la forma de la urna esto incluirá, o no, el orden de ingreso, de las bolitas, a la urna. Por ejemplo, si tenemos urnas como los envases de las pelotas de tenis, el orden será distinguible (¡la primera en entrar es la última en salir!), pero si es una urna grande respecto a las bolitas, este orden se pierde.

4.6 Permutaciones y Combinaciones

El caso más simple es considerar n bolitas distinguibles que se colocan en N casillas con, a lo sumo, un elemento por casilla. Ahora bien, para lograr que todas las bolitas queden en alguna casilla es necesario que $N \geq n$.

Nota: Ya hemos usado este procedimiento para resolver algunos ejercicios anteriores.

Por el Principio de Multiplicación el total de configuraciones distintas de urnas con bolitas es:

$$N \cdot (N-1) \cdot \dots \cdot (N-n+1) = \frac{N!}{(N-n)!}.$$

Colocar todas las bolitas equivale a asignar una urna a cada bolita. Hay N urnas disponibles para la primera bolita, $(N-1)$ casillas disponibles para la segunda bolita,$\dots$, $(N-(n-1))$ casillas disponibles para la n-ésima bolita.

Nota: Esto se conoce como el total de *n-permutaciones* de N elementos.

Cuando las bolitas ya no se distinguen entre sí (digamos que son todas negras). El total de configuraciones distintas disminuye y ahora es:

$$\binom{N}{n} := \frac{N!}{(N-n)! \cdot n!}.$$

Al no distinguir las bolitas, las configuraciones ahora sólo se diferencian por las urnas que tienen una bolita y las que no y es por eso que el número de configuraciones distintas disminuye. Como es habitual en matemática, la idea es aprovechar el trabajo ya realizado más que intentar todo de nuevo. Si a todas las bolitas de cada configuración con bolitas distinguibles las pintamos de modo que quedan todas iguales, obtenemos las configuraciones que estamos buscando. El total de configuraciones que dan una misma configuración nueva son aquellas en que las bolitas ocupan las mismas urnas pero en lugares diferentes. Como hay $n!$ maneras de repartirlas en lugares diferentes entonces el total de configuraciones distinguibles es ahora $\frac{1}{n!}$ de las configuraciones distinguibles originales. Otra manera de pensarlo es que se obtiene una misma configuración para todas las asignaciones de bolitas en que fijan las urnas que las reciban y luego se lanzan las bolitas en todos los órdenes posibles.

Nota: Esto se conoce como el total de *n-combinaciones* de N elementos.

La diferencia entre combinación y permutación es una fuente de confusión habitual en combinatoria. En nuestro contexto la diferencia es simple y depende, sólo de si, son distinguibles las bolitas que lanzamos (ver figura 4.5) y, por tanto, siempre el número de permutaciones será $n!$ veces el número de combinaciones al considerar n bolitas en N casillas.

4.6.1 Aplicaciones de Combinación y Permutación

Vamos a explicar en detalle algunas aplicaciones importantes de estas dos fórmulas.

Aplicación 4. *El total de subconjuntos de n elementos de un total de N elementos distintos, es:*

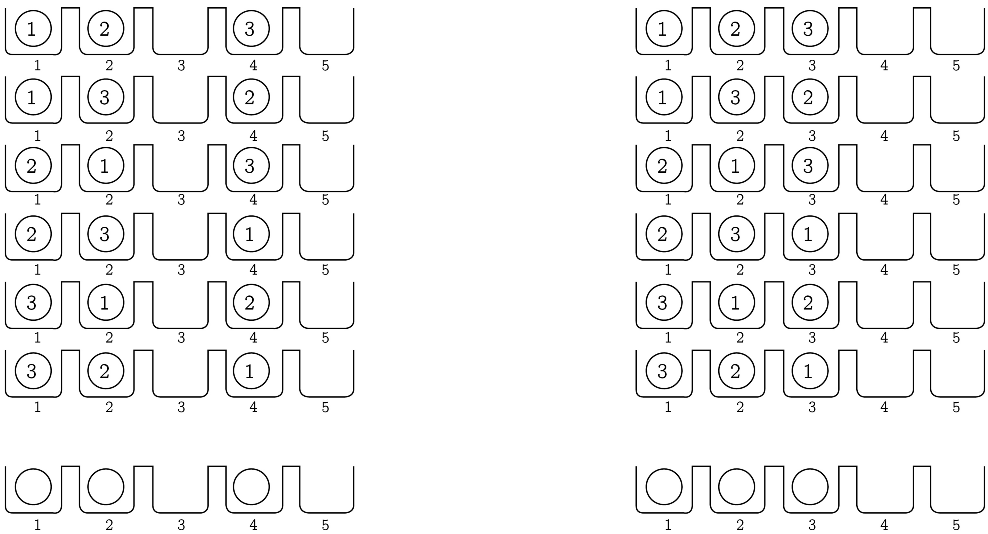


Figura 4.5. Caso N=4 y n=3

$$\binom{N}{n} := \frac{N!}{(N-n)!\,n!}$$

En efecto, los N elementos distintos son N casillas distintas y un subconjunto de n elementos se forma al llenar n urnas con n bolitas no distinguibles.

Aplicación 5 (Teorema del Binomio)**.** *Para cualquier par de números reales a y b y para cualquier natural N se tiene que:*

$$(a+b)^N = \sum_{n=0}^{N} \binom{N}{n} a^n b^{N-n}.$$

Por definición:

$$(a+b)^N = \overbrace{(a+b)\cdot \dots \cdot (a+b)}^{N}$$

Por la distributividad del producto sobre la suma, sabemos que $(a+b)^N$ es una suma de productos de N números. Esto se puede pensar como N bolsas con 2 elementos distinguibles a y b. Por aplicación directa del Principio de Multiplicación el total de sumandos del desarrollo es 2^N.

Pero esto no ha terminado, puesto que no todos estos elementos son distinguibles. De hecho, los productos distinguibles son sólo $N+1$ y de la forma $a^n \cdot b^{N-n}$ con n entre 0 y N. En otras palabras los elementos distinguibles quedan determinados al decidir cuántas a's van a tener.

Si hemos fijado la cantidad de veces que aparece el elemento a, digamos n, falta determinar en qué n posiciones van a quedar y esto corresponde al número de subconjuntos de n elementos de un conjunto de N que sabemos es $\binom{N}{n}$. Entonces, todos estos términos suman $\binom{N}{n} a^n \cdot b^{N-n}$. Al sumar sobre n obtenemos el Teorema.

Nota: Esta aplicación es la que da el nombre de *Coeficiente Binomial* a la cantidad $\binom{N}{n}$.

Usaremos algunas propiedades interesantes de los coeficientes binomiales que se pueden demostrar de manera combinatoria.

Proposición 4.1. *Dados* $N, n \in \mathbb{N}$ *y* $n \leq N$

1. *Si* $n \geq 1$, $\binom{N}{n-1} + \binom{N}{n} = \binom{N+1}{n}$.
2. $\binom{N}{0} + \binom{N}{1} + \cdots + \binom{N}{N} = 2^N$

Demostración. (1) La cantidad $\binom{N+1}{n}$ cuenta el total de subconjuntos de n elementos de un conjunto de $N+1$ elementos. Estos subconjuntos se pueden separar en dos grupos disjuntos, aquellos que contienen a un elemento en particular y el resto que no. Por definición, los que no contienen a este elemento son exactamente $\binom{N}{n}$ y los que sí lo contienen corresponden a escoger un subconjunto de $n-1$ elementos de N y agregarle ese elemento, lo que da un total de $\binom{N}{n-1}$. Por el principio de suma se concluye.

(2) Esta suma corresponde a la suma de todos los posibles subconjuntos de un conjunto de N elementos que ya vimos era 2^N. □

La propiedad (1) de la proposición 4.1 permite construir los Coeficientes Binomiales en secuencia, generando lo que se conoce como el *Triángulo de Pascal* (ver Figura 4.6).

4.7 Ejercicios de Permutación y Combinación

Como en secciones anteriores vamos a dar una lista de ejercicios y resolveremos algunos de ellos detalladamente.

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & \\
 & & & 1 & & 3 & & 3 & & 1 & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
$$

FIGURA 4.6. Triángulo de Pascal

1. ¿Cuántos comités distintos de 5 personas compuesto de 3 hombres y 2 mujeres se pueden formar con un total de 7 hombres y 5 mujeres?

   Respuesta: De los 7 hombres hay $\binom{7}{3}$ grupos de 3 hombres y de las 5 mujeres hay $\binom{5}{2}$ grupos de 2 mujeres. Por el Principio de Multiplicación el comité puede escogerse de $\binom{7}{3}\binom{5}{2} = 350$ maneras.
2. Todos los años, se selecciona una delegación de 4 estudiantes de un colegio para asistir a la asamblea anual de la Asociación de Estudiantes.

   a) ¿De cuántas maneras puede escogerse la delegación si hay 12 estudiantes elegibles?
   b) De todos los estudiantes se tiene dos hermanos que no están dispuestos a estar juntos en la delegación, ¿de cuántas maneras puede escogerse ahora la delegación?
   c) Ahora se tiene dos hermanos que no están dispuestos a asistir el uno sin el otro, ¿de cuántas maneras puede escogerse ahora la delegación?

   Respuesta: Una delegación corresponde a un subconjunto de 4 elementos.
   a) El total de delegaciones de 4 estudiantes de un grupo de 12 es $\binom{12}{4} = 495$.
   b) El total de delegaciones que no incluye a ninguno de los dos hermanos es $\binom{10}{4} = 210$. El total de delegaciones que incluye a un hermano, pero no al otro, son $\binom{10}{3} = 120$, puesto que se debe seleccionar sólo a tres miembros. Como son dos hermanos el total de delegaciones que incluye a uno de los dos, pero no los dos juntos es $2\binom{10}{3} = 240$. Por el principio de suma, entonces, el total de delegaciones es $210 + 240 = 450$.
   c) El total de delegaciones que no incluye a ninguno de los dos es $\binom{10}{4} = 210$ y el total de delegaciones que los incluye a ambos es $\binom{10}{2} = 45$. Por el principio de suma, la delegación puede escogerse entonces de $210 + 45 = 255$ maneras.

3. ¿Cuántas señales distinguibles de 8 banderas colocadas en una línea vertical pueden formarse con un conjunto de 4 banderas rojas, 3 blancas y una azul?

   Respuesta: El total de señales distinguibles si todas las banderas fuesen distinguibles es 8!. Dado que hay 4 banderas rojas que no se diferencian entre sí, todas estas maneras se pueden agrupar en grupos de 4! señales que no se distinguen. Lo mismo ocurre con el grupo de 3 banderas blancas por lo que el total de señales diferentes será entonces $\frac{8!}{4!3!} = 280$.

4. ¿De cuántas maneras distintas se puede acomodar una reunión de 7 personas?
   a) En una fila de 7 sillas
   b) Alrededor de una mesa redonda.
   Observación: Dos configuraciones de personas sentadas en una mesa redonda son distintas si alguna persona tiene un vecino distinto a la derecha o la izquierda.

   Respuesta:
   a) Las siete personas pueden distribuirse en una fila de 7! maneras distintas.
   b) Sin embargo, si están en una mesa redonda al rotar la mesa no se obtiene una configuración diferente puesto que todas las personas tienen los mismos vecinos a la izquierda y a la derecha. Por lo tanto la cuenta se debe hacer después de que sentamos a la primera persona. Es decir, las otras seis personas pueden acomodarse de 6! maneras diferentes alrededor de la mesa.
   Nota: Esto se conoce como *permutación circular*.
5. Un estudiante tiene que contestar 8 de 10 preguntas en un examen.
   a) ¿Cuántas grupos diferentes de 8 preguntas puede seleccionar?
   b) ¿Cuántas posibilidades tiene ahora, si las tres primeras preguntas son obligatorias?
   c) ¿Cuántas posibilidades tiene si hay que contestar 4 de las 5 primeras preguntas?
   Respuesta:
   a) El total de subconjuntos de 8 preguntas de un total de 10 es $\binom{10}{8} = 45$.
   b) Si contesta las 3 primeras preguntas entonces puede escoger las otras 5 de las 7 últimas preguntas lo que da un total de $\binom{7}{5} = 21$.
   c) Si contesta las 5 primeras preguntas entonces puede escoger las otras 3 de las 5 últimas lo que da $\binom{5}{3} = 10$. La otra posibilidad es que seleccione 4 de las 5 primeras preguntas y las otras 4 de las 5 últimas. El total de maneras de seleccionar 4 preguntas de 5 es $\binom{5}{4} = 5$. Por el Principio de Multiplicación puede escoger las 8 preguntas de 25 maneras. Para concluir, por el Principio de suma existen $10 + 25 = 35$ posibilidades a responder.
6. Un estudiante tiene 6 libros de matemática y 4 de física ordenados en un mismo estante. Hallar la probabilidad de que 3 libros determinados de matemática estén juntos.

Respuesta: El total de maneras de ordenar los libros es 10!. Los tres libros de matemáticas pueden ordenarse de 3! maneras. El resto de los 7 libros pueden ordenarse de 7! formas. Entonces, la probabilidad pedida será de $\frac{1}{120}$.

7. Una caja contiene 8 bolas rojas, 3 blancas y 9 azules; si se extraen 3 bolas aleatoriamente de la caja. Determinar la probabilidad de que:
   a) las 3 bolas sean rojas
   b) las 3 bolas sean blancas
   c) 2 sean rojas y 1 blanca
   d) al menos 1 sea blanca
   e) se extraiga una de cada color
   f) las bolas sean extraídas en el orden rojo, blanco, azul.
   a) Al considerar favorables sobre posibles como probabilidad, la respuesta pedida es el número de grupos de 3 bolas que se pueden tomar de un total de 8 bolas rojas, dividido por el número de grupos de 3 bolas que se pueden formar de un total de 20 bolas. Es decir, $\frac{\binom{8}{3}}{\binom{20}{3}} = 14285$.
   b) Como en el caso anterior esta probabilidad es $\frac{\binom{3}{3}}{\binom{20}{3}} = 11140$
   c) Una vez más la probabilidad será $\frac{\binom{8}{2}\cdot\binom{3}{1}}{\binom{20}{3}} = \frac{7}{95}$
   d) La probabilidad de no obtener ninguna bola blanca es $\frac{\binom{17}{3}}{\binom{20}{3}} = \frac{34}{57}$. La probabilidad de obtener al menos 1 blanca es complementaria, por lo tanto, será $1-\frac{34}{57} = \frac{23}{57}$
   e) La probabilidad de que aparezca una de cada color es $\frac{\binom{8}{1}\cdot\binom{3}{1}\cdot\binom{9}{1}}{\binom{20}{3}} = \frac{18}{95}$
   f) Usando lo anterior, la probabilidad de extraer las bolas en orden rojo, blanco y azul es $\frac{1}{3!}\cdot\frac{18}{95} = \frac{3}{95}$.
8. Se reparten 52 cartas de una baraja corriente en partes iguales a cuatro personas que llamaremos $Norte(N)$, $Sur(S)$, $Este(E)$ y $Oeste(O)$.
   a) Si S no tiene ases, hallar la probabilidad de que su compañero N tenga exactamente 2 ases.
   b) Si N y S juntos tienen 9 corazones, hallar la probabilidad de que E y O tengan cada uno dos corazones.

   Respuesta:
   a) Para cualquiera de los 4 jugadores hay $\binom{39}{13}$ grupos distintos de cartas que pueden recibir. Si S no tiene ases entonces hay 39 cartas, que necesariamente incluyen los 4 ases, repartidas entre N, E y O. Hay $\binom{4}{2}$ grupos de 2 ases que N puede recibir de los 4 ases y $\binom{35}{11}$ grupos de 11 cartas de las 35 cartas que no son ases. La probabilidad pedida entonces será:

$$\frac{\binom{4}{2}\cdot\binom{35}{11}}{3913} = \frac{650}{2109}.$$

b) Hay 26 cartas, que contienen exactamente 4 corazones, repartidas entre E y O. Hay $\binom{26}{13}$ grupos distintos de 13 cartas que E puede recibir que, además, fija el grupo de cartas que recibirá O. Hay $\binom{4}{2}$ pares de corazones que E puede recibir de los 4 posibles y $\binom{22}{11}$ grupos de 11 cartas de las 22 que no son corazones. Entonces, la probabilidad pedida es:

$$\frac{\binom{4}{2}\cdot\binom{22}{11}}{\binom{26}{13}} = \frac{234}{575}.$$

9. Muestre que el número de permutaciones de la secuencia *abcd*, que no incluyen *ab*,*bc* ni *cd* son 11. Verifique escribiendo todas las permutaciones.

10. Muestre que el número de permutaciones de la secuencia *abcd* en que a y b no quedan contiguas y b y c tampoco son:

$$4! - \binom{2}{1}\cdot 2\cdot 3! + \binom{2}{2}\cdot 2^2 = 8.$$

11. De un grupo de 6 hombres y 4 mujeres se forma un comité de 3 personas.

    a) ¿Cuántos comités distintos se pueden formar?

    b) ¿Cuántos comités distintos con un sólo hombre se pueden formar?

    c) ¿Cuántos comités distintos con al menos un hombre se pueden formar?

12. En una asamblea de n estudiantes se discute sobre 3 métodos alternativos de escoger un comité de $k \leq n$ estudiantes con un presidente. Un estudiante propone seleccionar el comité de k personas y dentro de él seleccionar al presidente. Otro sugiere seleccionar un grupo de $k-1$ personas y escoger el presidente entre los restantes. Finalmente otro estudiante sugiere seleccionar primero al presidente de entre todos los estudiantes y luego de los restantes escoger los $k-1$ miembros. Calcule en cada caso el número de comités distintos que se pueden formar y compare, ¿vale la pena discutir sobre cual método utilizar?

4.7.1 Bolitas semidistinguibles

Pasemos a un caso intermedio entre distinguir todas las bolitas y no distinguir a ninguna. Consideremos el caso en que las bolitas se pueden agrupar en dos tipos (digamos blancas y negras).

Aplicación 6. *De una urna con N bolitas se extraen n de ellas. Si del total de bolitas $m(\leq N)$ son blancas y las $N-m$ restantes son negras. El total de maneras distintas de extraer exactamente $r(\leq m)$ bolitas blancas es:*

$$\binom{m}{r} \cdot \binom{N-m}{n-r}.$$

En esta aplicación *no estamos lanzando bolitas sobre urnas*. Por el contrario, estamos extrayendo n bolitas de una urna con un total de N bolitas, lo que equivale a seleccionar un subconjunto de n elementos de un total de N. En apartados anteriores ya vimos cómo se podía entender en el contexto de lanzar bolitas sobre urnas.

Para obtener exactamente r bolitas blancas al extraer n bolitas, es necesario que las $n-r$ bolitas restantes sean todas negras. Es necesario, entonces, que $\max\{0, n-(N-m)\} \leq r \leq \min\{n, m\}$.

Consideremos un subconjunto de n bolitas que tiene r bolitas blancas y $n-r$ bolitas negras. El subconjunto de r bolitas blancas es uno de los subconjuntos de r bolitas de un total de m bolitas blancas. El subconjunto de $n-r$ bolitas negras es uno de los subconjuntos distintos de $n-r$ bolitas de un total de $N-m$ bolitas negras. Aplicando el Principio de Multiplicación esto prueba el resultado.

Esto es la cantidad de subconjuntos distintos de n bolitas distinguibles de las cuales hay r blancas y $n-r$ negras; esta cantidad es distinta si consideramos que sólo se distingue el color de la bolita. Al "borrar" los números se produce una distorsión porque las configuraciones que quedan iguales no tienen la misma cantidad de elementos (ver figura 4.7).

La diferencia entre considerar bolas distinguibles o no, para realizar un cálculo de probabilidades, nos lleva en este ejemplo a una paradoja. Por ejemplo, en el caso de la figura 4.7, al calcular la probabilidad de obtener una bola blanca y una negra, por un lado llegamos a la conclusión que la probabilidad es $\frac{6}{10}$ y por otro que la probabilidad es $\frac{1}{3}$.

Esto se produce por la ya mencionada Falacia de Equiprobabilidad, puesto que en el segundo caso no es correcto, dadas las condiciones del experimento, asumir que la probabilidad se calcula como favorables sobre posibles.

Una aplicación "real" de lo anterior se utiliza para estimar poblaciones que son demasiado grandes para realizar un método exhaustivo (contarlos a todos).

Ejemplo 1. *Para estimar la población N de salmones en un lago, se toma una muestra de m salmones que se marcan de alguna manera y se devuelven al lago. Se espera un tiempo razonable antes de tomar una segunda muestra de n salmones. Se cuentan los ejemplares marcados, digamos k. Una estimación razonable de N sería considerar el valor de N que maximiza la probabilidad de que haya k salmones en el lago. Este*

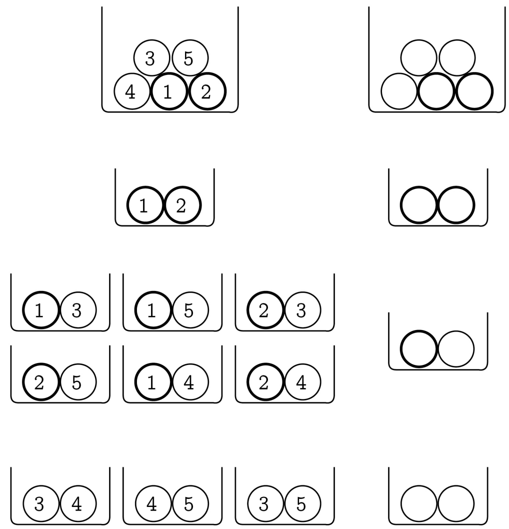


FIGURA 4.7. Caso N=5 n=3 y r=2

tipo de razonamientos se conoce en estadística como estimador máximo verosímil.

Utilizando el razonamiento anterior la probabilidad de obtener k salmones marcados de una muestra de n sabiendo que hay N especímenes es:

$$\frac{\binom{m}{k} \cdot \binom{N-m}{n-k}}{\binom{N}{n}}.$$

¿Cómo se encuentra valor de N que maximiza este suceso? La respuesta a este tipo de preguntas se ven en cursos de *Inferencia Estadística.*

La generalización del problema de los tipos es la siguiente:

Aplicación 7. *Se tienen N bolitas, m_1 de tipo 1, m_2 de tipo 2,$\cdots$, m_n de tipo n con $\sum_{i=1}^{n} m_i = N$. Se lanzan las N bolitas sobre N urnas con a lo más una bolita por urnas. Si sólo se distingue el tipo de las bolitas, el total de configuraciones distiguibles es:*

$$\frac{N!}{m_1!m_2!\cdot\dots\cdot m_n!}$$

En efecto hay $N!$ configuraciones distintas de N bolitas distinguibles en N urnas. Sin embargo, dado que no se distinguen todas las bolitas entre sí hay varias configuraciones que son iguales con respecto a los tipos. Más precisamente, en las urnas que tienen bolitas del mismo tipo, es posible intercambiarlas, sin cambiar la configuración. Esto se puede hacer de $m_1!$ para el primer tipo de elementos, de $m_2!$ maneras para el segundo tipo,..., de $m_n!$ maneras para el último tipo. Por lo que hay $m_1!\cdot m_2!\cdot\cdots\cdot m_n!$ formas de reordenar sin cambiar la configuración de los tipos.

En consecuencia, $m_1!\cdot m_2!\cdot\cdots\cdot m_n!$ configuraciones de bolitas distinguibles dan una sola configuración distinguible con respecto a los tipos. Dado que el total de configuraciones era $N!$ se concluye.

Una aplicación directa de esto permite demostrar el *Teorema del Multinomio* que dice que dados $a_1,\ldots,a_n$ números reales se tendrá que:

$$(a_1+\cdots+a_n)^N = \sum_{m_1=0}^{n} \sum_{m_2=0}^{N-m_1} \cdots \sum_{m_n=0}^{N-\sum_{i=0}^{n-1} m_i} \frac{N!}{m_1!\cdot m_2!\cdot\cdots\cdot m_n!} a_1^{m_1}\cdot\cdots\cdot a_n^{m_n}.$$

La demostración sigue las mismas líneas que en el caso del Teorema del Binomio.

4.8 Urnas con más de 1 Bolita

Consideramos ahora un caso más complejo en que se lanzan N bolitas sobre n urnas permitiendo más de una bolita por urna. En este caso los roles de n y N están invertidos con respecto a la secciones anteriores, puesto que en algunos casos necesitaremos más bolitas que urnas.

4.8.1 Grupos Permitiendo Casillas Vacías

El total de configuraciones distintas de urnas con bolitas distinguibles permitiendo algunas casillas vacías es:

$$\frac{(N+n-1)!}{(n-1)!}.$$

En la figura 4.8 se muestran las configuraciones posibles con 2 bolitas y 3 casillas. Si las bolitas ahora no son distinguibles, el total de configuraciones distintas aceptando casillas vacías es:

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

En el caso de la figura 4.8 las configuraciones distintas se reducen ahora a las mostradas en la figura 4.9.

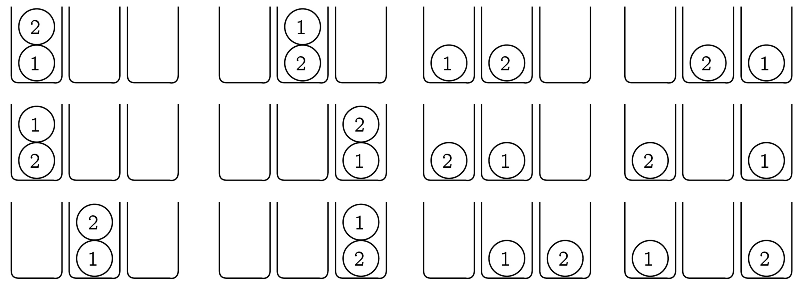


FIGURA 4.8. Caso N=2, n=3 (Total=12)

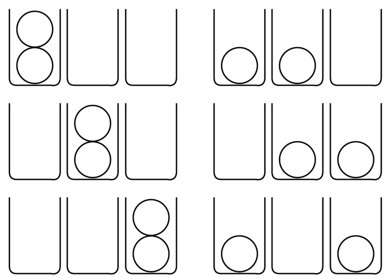


FIGURA 4.9. Caso N=2, n=3 (Total=6)

Estas fórmulas se pueden demostrar usando las técnicas que hemos visto hasta ahora.

Dado un orden posible de las N bolitas, se agregan $n-1$ "marcas" y se consideran todas las configuraciones posibles (que incluyen todos los órdenes de los elementos) que son $(N+n-1)!$. Sin embargo, hay $n-1$ objetos iguales (las marcas) por lo que el total de configuraciones distintas es:

$$\frac{(N+n-1)!}{(n-1)!}.$$

Para ilustrar el razonamiento consideremos los posibles resultados para el ejemplo de la figura 4.8. Con esta nueva notación si denotamos las marcas por el símbolo $|$, entonces, esto se reescribe como:

$$\begin{array}{cccc} 12|| & |21| & 1|2| & |2|1 \\ 21|| & ||12 & 2|1| & 2||1 \\ |12| & ||21 & |1|2 & 1||2 \end{array}$$

Si no se distinguen las bolitas entre sí entonces el total de configuraciones distintas disminuye en un factor de $N!$ (como al pasar de combinaciones a permutaciones) y, por lo tanto, da:

$$\frac{(N+n-1)!}{(n-1)!\cdot N!} = \binom{N+n-1}{n-1}.$$

Una vez más para ilustrar el razonamiento correspondiente al ejemplo de la figura 4.9 con esta notación notando las marcas por $|$ y las bolitas por $\bigcirc$ los resultados se reescriben como:

$$\begin{array}{ll} \bigcirc\bigcirc|| & \bigcirc|\bigcirc| \\ |\bigcirc\bigcirc| & |\bigcirc|\bigcirc \\ ||\bigcirc\bigcirc & \bigcirc||\bigcirc \end{array}$$

4.8.2 Grupos sin Casillas Vacías

Si sólo se consideran las configuraciones en que todas las casillas tienen alguna bolita será necesario que $N \geq n$. En ese caso, el total de configuraciones será:

$$N!\binom{N-1}{n-1}.$$

En la parte izquierda de la figura 4.10 se muestran las configuraciones para el caso $N = 4$ y $n = 3$. Al considerar que las bolitas no son distinguibles entre sí el total de configuraciones en que hay al menos una bolita por urna disminuye y es:

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

En la parte derecha de la figura 4.10 se muestran las configuraciones distintas para el mismo caso, pero sin distinguir las bolitas (en vez de 72 sólo hay 3!!).

En este caso, la construcción que utilizamos anteriormente debe ser un poco diferente. Dado un orden ($N!$ posibles) de las N bolitas. Repartir las bolitas en n urnas sin dejar ninguna vacía equivale a repartir $n-1$ "marcas" como anteriormente, pero ahora sólo entre los $N-1$ espacios entre las bolitas. Esto se puede hacer de $\binom{N-1}{n-1}$ maneras.

Por lo tanto hay, $N!\binom{N-1}{n-1}$ configuraciones distintas. Como en los casos anteriores para ilustrar el razonamiento los posibles resultados para el ejemplo considerado en la figura 4.10 notando las marcas por el símbolo $|$:

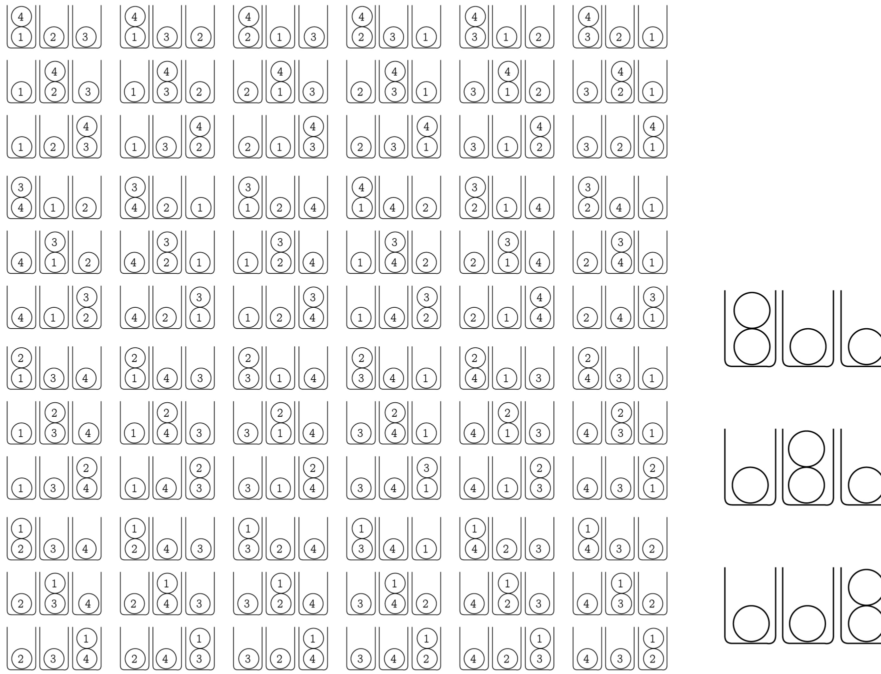


FIGURA 4.10. Caso N=4, n=3 (Total=72)(Total=3)

14|2|3	14|3|2	24|1|3	24|3|1	34|1|2	34|2|1
1|24|3	1|34|2	2|14|3	2|34|1	3|14|2	3|24|1
1|2|34	1|3|24	2|1|34	2|3|14	3|1|24	3|2|31
43|1|2	43|2|3	13|2|4	13|4|2	23|1|4	23|4|1
4|13|2	4|23|3	1|23|4	1|14|2	2|13|4	2|43|1
4|1|23	4|2|33	1|2|43	1|4|23	2|1|43	2|4|13
12|3|4	12|4|3	32|1|4	32|4|1	42|1|3	42|3|1
1|32|4	1|42|3	3|12|4	3|42|1	4|12|3	4|32|1
1|3|42	1|4|43	3|1|42	3|4|12	4|1|32	4|3|12
21|3|4	21|4|3	31|2|4	31|4|2	41|2|3	41|3|2
2|31|4	2|41|3	3|21|4	3|41|2	4|21|3	4|31|2
2|3|41	2|4|31	3|2|41	3|4|21	4|2|31	4|3|21

Si las bolitas no se distinguen entre sí entonces el total de configuraciones disminuye en un factor de $N!$, lo que da $\binom{N-1}{n-1}$. Finalmente, para el ejemplo considerado en la figura 4.10 notando las marcas y las bolitas como en el caso anterior obtenemos las posibilidades

○○|○|○ ○|○○|○ ○|○|○○

Otra manera de deducir las fórmulas para este caso, es considerar, que se saca un grupo de n bolitas del total de N y se coloca una en cada casilla y ahora se consideran todas las configuraciones de $N - n$ bolitas en n casillas permitiendo, de este modo, casillas vacías. ¿Se obtienen los mismos resultados?.

4.9 El Problema del Coleccionista de Cupones

Veamos una aplicación no directa del resultado anterior que se conoce como el "Coupon Collector Problem" o *Problema del Coleccionista de Cupones*.

Aplicación 8. *Un coleccionista de cupones va a comprar n cupones de un total de N, ¿cuál es el total de configuraciones distintas (cuántos cupones de cada tipo) que puede obtener el coleccionista?, son todas las configuraciones igualmente probables?*

En un contexto general dada una infinidad de elementos de N tipos distintos el total de configuraciones distintas (al reconocer por tipo) de n de estos elementos es:

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

Si se intenta resolver pensando en cada tipo como una bolita, la cuenta es imposible de hacer con los resultados que tenemos. Sin embargo, podemos considerar los tipos como urnas y cada elemento será una bolita que deberá caer en alguna casilla (el tipo que le corresponde). Por lo tanto, por el resultado anterior se tendrá que el total de configuraciones distintas de n bolitas no distinguibles que se reparten en N casillas es:

$$\binom{n+N-1}{N-1} = \binom{n+N-1}{n}.$$

Esto muestra que las configuraciones no son todas igualmente probables.

4.10 Ejemplos Propuestos

1. ¿De cuántas maneras 3 americanos, 4 franceses, 4 daneses y 2 italianos pueden sentarse en una fila, de modo que aquellos de la misma nacionalidad se sienten contiguos?. Resolver el mismo problema si se sientan en una mesa redonda.

   Respuesta: Existen 4! órdenes posibles para que las 4 nacionalidades se formen en una fila. Hay 3! maneras de sentar a los americanos contiguos, 4! maneras de sentar a los 4 franceses contiguos, 2! maneras de sentar a los italianos contiguos. Por el Principio de Multiplicación el total de maneras de sentarlos es $4! \cdot 3! \cdot 4! \cdot 4! \cdot 2! = 165{,}888$.

Como ya se había comentado en problemas anteriores, al considerar una mesa redonda del total de 4! órdenes posibles para las nacionalidades es, en realidad, sólo 3! puesto que al rotar la mesa los vecinos no cambian. Hay 3! maneras de sentar a los 3 americanos contiguos, 4! maneras de sentar a los 4 franceses contiguos, 4! maneras de sentar a los 4 daneses y 2! maneras de sentar a los 2 italianos. Por el Principio de Multiplicación esto da $3!\cdot 3!\cdot 4!\cdot 4!\cdot 2! = 41,472$ órdenes posibles.

2. ¿Cuántas señales diferentes, cada una de 6 banderas colgadas en una línea vertical, pueden formarse con 4 banderas rojas idénticas y 2 azules idénticas?
   Respuesta: $\frac{6!}{4!2!} = 15$.
3. ¿Cuántas permutaciones distintas pueden formarse con todas las letras de cada una de las palabras (i) tema, (ii) campana, (iii) estadísticas?
   Respuesta: (i) 24, (ii) 840, (iii) $\frac{12!}{3!2!2!2!}$.
4. Cuatro libros diferentes de matemática, seis diferentes de física y dos diferentes de química se colocan en un estante, ¿de cuántas maneras distintas es posible ordenarlos si: (a) los libros de cada asignatura deben estar todos juntos, (b) solamente los libros de matemática deben estar juntos?
   Respuesta: (a) 4!6!2!3!=207.360, (b) 9!4!=8.709.120.
5. Se ordenan en una fila 5 bolas rojas, 2 bolas blancas y 3 bolas azules. Si las bolas de igual color no se distinguen entre sí, ¿De cuántas formas posibles pueden ordenarse?
   Respuesta: $\frac{10!}{5!2!3!}$.
6. De un total de 5 matemáticos y 7 físicos, se forma un comité de 2 matemáticos y 3 físicos, ¿de cuántas formas puede formarse el comité, si (a) puede pertenecer a él cualquier matemático y físico; (b) un físico determinado debe pertenecer al comité, (c) dos matemáticos determinados no pueden estar en el comité?
   Respuesta: (a) 350, (b) 150, (c) 105.
7. ¿Cuántas ensaladas pueden prepararse con lechuga, escarola, endibia, berro y achicoria?
   Respuesta: 31.
8. Se extraen 5 cartas de una baraja de 52 cartas. Hallar la probabilidad de extraer: (a) 4 ases; (b) 4 ases y un rey; (c) 3 diez y 2 jotas; (d) un 9, 10, jota, reina, rey en cualquier orden; (e) 3 de un palo y 2 de otro; (f) al menos un as.
   Respuesta: (a) $\frac{1}{54,145}$, (b) $\frac{1}{649,740}$, (c) $\frac{1}{108,290}$, (d) $\frac{64}{162,435}$, (e) $\frac{429}{4,165}$, (f) $\frac{18,472}{54,145}$.
9. Determinar la probabilidad de aparición de tres veces 6 en 5 lanzamientos de un dado equilibrado.
   Respuesta: $\frac{125}{3,888}$.
10. A un jugador le reparten 5 cartas, una tras otra, de una baraja corriente de 52 cartas, Cuál es la probabilidad de que todas sean espadas?
    Respuesta: $\frac{33}{66,640}$
11. Se lanzan tres monedas corrientes. Hallar la probabilidad de que sean todas caras sí: (i) la primera de las monedas es cara; (ii) una de las monedas es cara.
    Respuesta: (i) $\frac{1}{4}$, (ii) $\frac{1}{7}$.

12. ¿De cuántas maneras puede un profesor escoger uno o más estudiantes de seis elegibles?
    Respuesta: 63.
13. Se dispone de 3 tipos de pruebas con 4 copias cada una. ¿De cuantas maneras se pueden repartir entre 12 estudiantes?.
    Respuesta: $\binom{12}{4}\binom{8}{4} = 34{,}650$.
14. ¿De cuántas maneras 12 estudiantes pueden repartirse en 3 equipos, de modo que cada equipo esté formado por 4 estudiantes?
    Respuesta: $\binom{11}{3}\binom{7}{3} = 5{,}775$.
15. Calcular el total de configuraciones distintas que se pueden obtener, al repartir 5 anillos en 4 dedos de una mano (no se considera el pulgar), cuando los anillos son distintos y cuando no lo son.

16. Se tienen bolitas rojas, blancas y negras. Calcule el total de configuraciones distintas de 5 bolitas.

17. Una comisión de 10 personas está formada por 2 Ingleses, 2 Escoceses, 2 Japoneses y el resto de otras nacionalidades (todas distintas), ¿de cuantas maneras pueden ponerse en fila sin que haya dos de la misma nacionalidad contiguos?, ¿de cuantas maneras distintas se pueden sentar en una mesa redonda?

18. ¿Con cuántos mástiles y banderas es posible componer 126 mensajes si las banderas no son distinguibles entre sí?. Es decir, un mensaje es una cierta cantidad de banderas en cada mástil.

19. En el código Morse se emplean sólo dos signos: punto y raya, ¿cuantas palabras distintas de n signos se pueden obtener utilizando el código Morse?.

20. ¿Cuántos números mayores que 3000 y menores que 4000 se pueden formar con los dígitos 2,3,6 y 9 si:
    a) Cada dígito se ocupa una sola vez.
    b) Cada dígito se puede ocupar más de una vez.
21. ¿Cuántas palabras diferentes se pueden obtener con las letras de la palabra MANICOMIO.

22. Si se toman n puntos distintos de una circunferencia y se unen todos con todos, ¿cuál es el máximo número de intersecciones posibles entre los trazos formados?.

23. Pedro va a comprar 10 frutas entre naranjas, manzanas y plátanos al supermercado, ¿cuantas posibilidades diferentes tiene Pedro para comprar?.

24. Se saca dos veces una bola de una urna con N bolitas numeradas de 1 a N (siempre se devuelven las bolitas). Calcule la probabilidad de que el número de la primera bola sea mayor estricto que el número de la segunda bola extraída.
    - ¿Cómo cambia esto si las bolitas no se devuelven?.
    - ¿Cuál es la probabilidad de que estén ordenadas de mayor a menor?.
    - ¿Cómo cambia la probabilidad que estén ordenadas si las bolitas no se devuelven?
25. En un sorteo de lotería se extraen 15 números de 25 posibles y el ganador debe obtener 15 aciertos. Cada boleto contiene 15 números y cuesta \$500.
    a) ¿Cuál es la probabilidad de ganar con un boleto?
    b) En el ejemplo anterior se ofrece un boleto especial que contiene 2 números comodines que pueden reemplazar cualquiera de los 15. Suponiendo que uno pudiese comprar fácilmente cualquier cantidad de boletos, ¿cuanto debería costar ese boleto especial?
26. Un cartero tiene N cartas que entregar en N destinos distintos.
    a) Suponiendo que el cartero pasó a celebrar las fiestas patrias antes de hacer su recorrido y, en consecuencia, entrega cada carta en cualquiera de los N destinos posibles y ni siquiera recuerda los destinos que ya ha visitado. Calcule la probabilidad de que k cartas lleguen a su destino con k en $\{0, \ldots, N\}$.
    b) Si el cartero estaba sobrio y entregó una sola carta en cada destino, pero algún gracioso desordenó las cartas. Calcule la probabilidad de que ninguna carta llegue a su destino.

4.11 Combinatoria y Funciones

La manera formalmente correcta de hacer combinatoria, es considerar como en el caso del cardinal, un conteo a base de funciones sobreyectivas o inyectivas. Esto es claramente menos didáctico que considerar bolitas y urnas. Pero por desgracia, no para todos los problemas es posible utilizar el esquema de bolitas y urnas (al menos de manera útil y didáctica).

Es claro que una realización de lanzar n bolitas distinguibles sobre N casillas define una única función $f : \{1, \ldots, n\} \to \{1, \ldots, N\}$ siendo $f(n)$ el número de la casilla donde se colocó la bolita n. Sin embargo, es posible que dos realizaciones distintas den la misma función por lo que esta relación no es siempre uno a uno.

Esto, porque la función no considera en que orden se lanzaron las bolitas. Por ejemplo, en la figura 4.8 las realizaciones que tienen dos bolitas en una casilla están contadas dos veces (primero la bola 1 y luego la bola 2 y viceversa) y las realizaciones que tienen una bolita por casilla están contadas una sóla vez.

La función será *inyectiva cuando sólo se acepte una bolita por casilla* y será *sobreyectiva si, ninguna urna queda sin bolitas*. En el contexto de bolitas y urnas esto es equivalente, sólo, en el caso en que se considera una bolita por urna. Veamos algunos ejemplos que se plantean en ambos contextos:

1. El total de funciones del conjunto $\{1, \ldots, n\}$ en el conjunto $\{1, \ldots, N\}$ es N^n.

   Es decir, lanzar n bolitas sobre N casillas permitiendo varias bolitas en las casillas.

2. Si $n \leq N$ el total de funciones inyectivas del conjunto $\{1, \ldots, n\}$ en el conjunto $\{1, \ldots, N\}$ es $\frac{N!}{(N-n)!}$.
   Es decir, lanzar n bolitas sobre N casillas con a lo más una bolita por casilla.

3. Si $N = n$ el total de funciones biyectivas del conjunto $\{1, \ldots, N\}$ en el conjunto $\{1, \ldots, N\}$ es $N!$.

   Es decir, lanzar N bolitas sobre N casillas con, a lo más , una bolita por casilla.

Con este nuevo formalismo algunos problemas de conteo se pueden plantear como determinar el número de *puntos fijos* de una función f del conjunto $\{1, \ldots, n\}$ en $\{1, \ldots, N\}$. Donde un punto fijo $i \in \{1, \ldots, N\}$ de la función f es aquel que satisface que $f(i) = i$.

Aplicación 9. *: Si $k \in \{0, \ldots, n\}$ entonces el total de funciones del conjunto $\{1, \ldots, n\}$ en el conjunto $\{1, \ldots, N\}$ que tienen k puntos fijos es:*

$$\binom{n}{k}(N-1)^{n-k}.$$

En efecto, para que una función f tenga k puntos fijos debe haber un subconjunto de k elementos de $\{1, \ldots, n\}$, donde se tienen los puntos fijos y además en los $n - k$ elementos restantes no se tendrá dicho punto. Hay $\binom{n}{k}$ subconjuntos distintos posibles para los puntos fijos. Una vez que se fijaron los elementos donde se alcanzarán dichos puntos hay una sóla posibilidad para la imagen de esos elementos y, para los elementos que no serán puntos fijos hay $N - 1$ posibilidades.

Nota: ¿Qué ocurre si sólo consideramos las funciones biyectivas de $\{1, \ldots, n\}$ en $\{1, \ldots, n\}$ con k puntos fijos?.

4.12 El Problema de los Prisioneros

Para finalizar esta sección vamos a plantear un problema, aparentemente, simple de resolver, que requiere de una solución bastante inusual para los procedimientos que hemos realizado hasta el momento. Esto introduce otra rama de técnicas combinatorias en las que se razona de manera "inductiva", o bien, "recursiva". Hasta el momento, todos los argumentos combinatorios que se han utilizado se pueden resumir en "sumar y multiplicar cuentas". Es posible, aunque parezca inútil, *contar de más para luego restar los elementos que se contaron más de una vez.*

Ejemplo 2. *Un grupo N soldados va a fusilar a n prisioneros con $N \geq n$. Cada uno de los soldados selecciona un prisionero al "azar" y dispara. Suponiendo que los prisioneros son un blanco fácil, ¿cuál es la probabilidad de que no sobreviva ninguno?.*

Nota: Es posible replantear este problema con amigos repartiendo regalos, pero mi experiencia muestra que este contexto particular hace que el problema no se olvide.

Se podría dar la siguiente solución al problema de los prisioneros. Cada bala es una bolita y cada prisionero es una urna. Por lo tanto, estamos lanzando las N bolitas en n urnas, aceptando más de una bolita por urna, permitiendo además, urnas sin bolitas sin reconocer las bolitas entre sí. En este caso el total de configuraciones posibles es:

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

El total de configuraciones que no dejan prisioneros vivos equivale al total de configuraciones con las condiciones anteriores que no dejan casillas sin bolitas, es decir:

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

Por lo tanto, la respuesta será:

$$\frac{\binom{N-1}{n-1}}{\binom{N+n-1}{n-1}}.$$

En el caso $N = n = 2$ se tendría probabilidad es $\frac{1}{\binom{3}{1}} = \frac{1}{3}$. Pero contando directamente, las posibles selecciones para los soldados son $(1,1)$, $(1,2)$, $(2,1)$, $(2,2)$, es decir, probabilidad $\frac{1}{2}$. Una vez más nos encontramos con una contradicción y una vez más se produce por la Falacia de Equiprobabilidad.

En efecto para $N = n = 2$ matar a ambos prisioneros se puede lograr de 2 maneras mientras que matar exactamente a uno, cualquiera sea, sólo se puede lograr de una manera.

La solución correcta es de hecho bastante más compleja de obtener.

Nota: Cada fusilamiento corresponde a seleccionar al azar una función de $\{1, \ldots, N\}$ en $\{1, \ldots, n\}$ y matarlos a todos corresponde a que la función sea sobreyectiva.

Como cada soldado selecciona un blanco de n posibles y hay N soldados, entonces por el principio de multiplicación hay n^N posibles resultados.

De los resultados posibles, hay $(n-1)^N$ maneras de que no muera el primer prisionero, por lo tanto, hay $n^N - (n-1)^N$ maneras de que el primer prisionero muera.

Las maneras en que el primer prisionero muere y el segundo no muere son, entonces:

$$(n-1)^N - (n-2)^N.$$

En efecto, basta eliminar la posibilidad de escoger el segundo y seguir el razonamiento anterior.

Si restamos a las maneras de matar al primer prisionero, aquellas en que no muere el segundo, obtenemos las maneras de matar al primer y al segundo prisionero.

Por lo tanto, las maneras de matar al primer y al segundo prisionero son:

$$n^N - 2(n-1)^N + (n-2)^N.$$

Con el mismo razonamiento las maneras de matar al primer y al segundo prisionero, sin matar al tercero, son:

$$(n-1)^N - 2(n-2)^N + (n-3)^N.$$

Una vez más, por el mismo razonamiento, las maneras de matar al primer, segundo y tercer prisionero son:

$$n^N - 3(n-1)^N + 3(n-2)^N - (n-3)^N.$$

Si se sigue el mismo proceso hasta agotar los prisioneros se concluye que las maneras de matarlos a todos son:

$$n^N - \binom{n}{1}(n-1)^N + \binom{n}{2}(n-2)^N + \cdots + (-1)^{n-1}\binom{n}{n-1}.$$

Por lo tanto, la probabilidad de que no sobreviva ninguno es:

$$\frac{n^N - \binom{n}{1}(n-1)^N + \binom{n}{2}(n-2)^N + \cdots + (-1)^{n-1}\binom{n}{n-1}}{n^N}.$$

Este ejemplo es una aplicación de una técnica que se conoce como el *Principio de Inclusión-Exclusión*.

4.13 Discos, Bolitas y Urnas reunidos

En esta sección vamos a conectar discos con urnas y con bolitas. Resumiendo lo que se ha estudiado en este capítulo, hasta el momento, podemos decir que introdujimos la combinatoria a través de bolitas y urnas y luego mostramos cómo todo esto se relacionaba con el estudio de funciones en un conjunto discreto. Ambos métodos tenían sus fortalezas y debilidades. Contando con ambos modelos y sabiendo relacionarlos se pueden aprovechar las fortalezas de cada uno de ellos y mejorar sus debilidades. Esta idea va más allá de las bolitas y las urnas y corresponde a una manera de razonar en matemática que podríamos llamar la “estrategia del mínimo esfuerzo con el máximo resultado”. La idea es aprovechar al máximo el trabajo ya realizado al tratar de resolver un problema nuevo. No se rehuye el trabajo, pero sí se intenta trabajar utilizando el mínimo de esfuerzo al largo plazo. Es decir, desarrollando métodos que permitan resolver problemas similares sin demasiado esfuerzo adicional.

Ahora consideramos que nuestras urnas ya tienen en su interior cierta cantidad conocida de bolitas y nuestro experimento consiste en extraer bolitas una a una. Se sabe que en todo momento cualquier bolita tiene igual probabilidad de ser extraída.

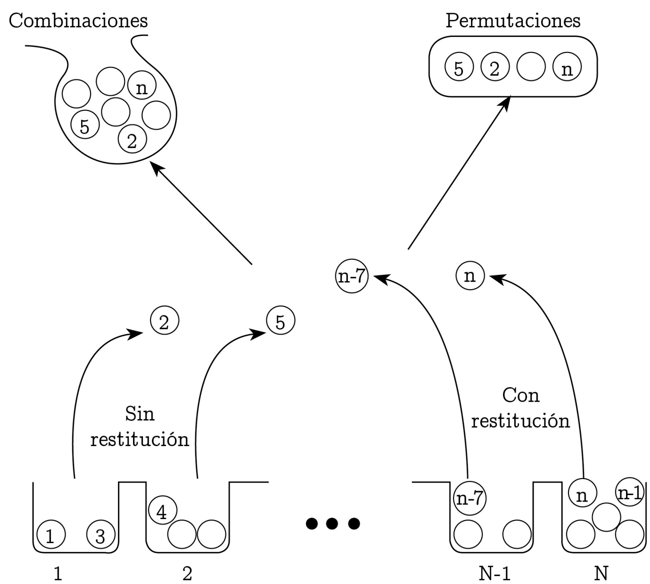


Figura 4.11. Extracción de Bolitas en Urnas

Existirán dos tipos distintos de extracción, las extracciones *con restitución* y las *sin restitución*. En el primer caso la bolita extraída se devuelve a la urna y en el segundo la bolita desaparece del sistema.

Nuestros resultados posibles serán secuencias de bolitas o grupos de bolitas dependiendo de si consideramos o no el orden en que fueron extraídas. Gráficamente una secuencia de bolitas quiere decir que las bolitas se guardan (o copian si hay restitución) en un tubo, de tal modo, que la última bolita extraída está al principio del tubo y la primera extraída está en el fondo del tubo. Por otro lado, un grupo de bolitas consiste simplemente en colocar las bolitas extraídas en un saco o bolsa (ver figura 4.11).

Comencemos con el caso simple de una urna y N bolitas distinguibles. Si extraemos $n(\leq N)$ bolitas, el total de secuencias distintas de bolitas es $\dfrac{N!}{(N-n)!}$ que coincide con el número de permutaciones distintas de n elementos de un total de N. Esto no es una coincidencia ya que corresponde a lanzar n bolitas sobre N urnas con, a lo sumo, una bolita por casilla.

En efecto, el primer elemento de la secuencia es el número de la urna donde cayó la bolita con el número 1, el segundo elemento de la secuencia es el número de la urna donde cayó la bolita 2, etc... La relación inversa es clara, dada una secuencia de bolitas: el primer número es la urna donde se coloca la bolita 1, el segundo número es la urna donde se coloca la bolita 2,etc...

Por ejemplo, las configuraciones de 4 urnas y 3 bolitas se traducen como:

$1-2-3$	$1-2-4$	$1-3-4$	$2-3-4$
$1-3-2$	$1-4-2$	$1-4-3$	$2-4-3$
$2-1-3$	$2-1-4$	$3-1-4$	$3-2-4$
$3-1-2$	$4-1-2$	$4-1-3$	$4-2-3$
$2-3-1$	$2-4-1$	$3-4-1$	$3-4-2$
$3-2-1$	$4-2-1$	$4-3-1$	$4-3-2$

Esto establece una biyección entre los dos conjuntos de configuraciones lo que muestra que tienen igual cardinal que en este caso es la cantidad de elementos.

Si ahora en el mismo experimento de urnas consideramos los grupos de bolitas distintos que podemos obtener, en vez de las secuencias, obtenemos $\binom{N}{n}$ que corresponde al número de subconjuntos de n elementos de un total de N.

Esto nos muestra una manera alternativa de interpretar la diferencia entre permutación y combinación. En las permutaciones consideramos el orden de extracción de las bolitas y en las combinaciones, sólo nos interesa el conjunto de bolitas obtenidas.

4.14 Extrayendo Bolas Semidistinguibles

Desarrollaremos algunos ejemplo particulares para ejemplificar las dificultades clásicas.

4.14.1 Ejemplo con 1 Urna

Veamos otro ejemplo que muestra esta manera de proceder. Consideremos el siguiente experimento: Se tiene una urna con 3 bolitas negras y 2 bolitas blancas y se extraen dos bolitas distintas de la urna.

En nuestra definición del experimento, para obtener dos bolitas de la urna estamos extrayendo las bolitas una por una sin restitución. De inmediato surge una duda físico-filosófica con respecto a lo que quiere decir extraer más de una bolita sin restitución. Físicamente, es diferente sacar bolitas una por una que sacarlas en puñados. Sin embargo, desde nuestro punto de vista no hay diferencia alguna, puesto que los resultados posibles son los mismos. Excepto, claro está, al considerar secuencias de bolitas y no grupos porque en ese caso no se podrá determinar el orden de extracción de las bolitas al sacarlas en puñados.

¿Cuál es la probabilidad de obtener dos bolitas blancas?, ¿la probabilidad de obtener dos bolitas negras?, ¿de obtener una bolita de cada color?. Volvamos a nuestra fórmula básica de favorables sobre posibles.

Podemos resolver el problema de al menos dos maneras distintas. Si estamos considerando el orden de extracción de las bolitas, entonces, el total de resultados posibles para dos bolitas es $5 \cdot 4 = 20$. El total de favorables para obtener dos bolitas blancas es $2 \cdot 1 = 2$ y el total de favorables para obtener dos bolitas negras es $3 \cdot 2 = 6$. Podríamos calcular el total de favorables para una bolita de cada color pero sabemos que debe ser $20 - 2 - 6 = 12$.

Por lo tanto, la probabilidad de obtener dos bolitas blancas es $\frac{2}{20} = \frac{1}{10}$, la probabilidad de obtener dos bolitas negras es $\frac{6}{20} = \frac{3}{10}$ y la probabilidad de obtener una bolita de cada color es $\frac{12}{20} = \frac{6}{10}$ (ver figura 4.12).

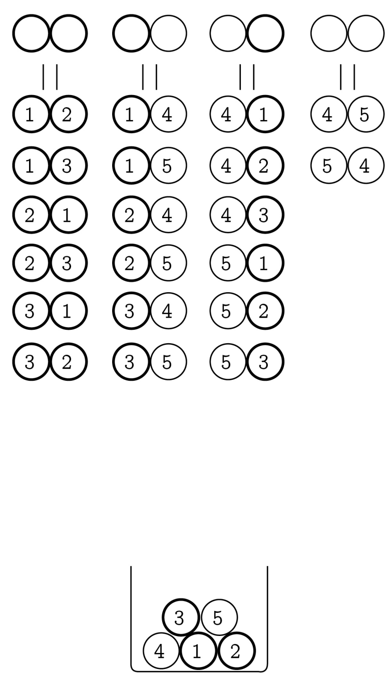


Figura 4.12. Extracción sin restitución

Si ahora no consideramos el orden de extracción, entonces el total de subconjuntos diferentes de 2 bolitas de un total de 5 es $\binom{5}{2} = \dfrac{5!}{3!2!} = 10$, que es la mitad de los resultados que consideramos al incluir el orden de extracción. El total de subconjuntos de dos bolitas blancas es $\binom{2}{2} = 1$ y el total de subconjuntos de dos bolitas negras es $\binom{3}{2} = 3$ y, por lo tanto el total de subconjuntos de bolitas una de cada color debe ser $10 - 1 - 3 = 6$. Al evaluar las probabilidades obtenemos exactamente los mismos resultados que antes. Esto ocurre porque todos los subconjuntos de dos bolitas son igualmente probables al suponer que todas las extracciones de dos bolitas considerando el orden son equiprobables.

4.14.2 Ejemplos con 3 Urnas

Consideremos ahora tres urnas. La primera urna es como la urna del ejemplo anterior, es decir, con 3 bolitas negras y 2 bolitas blancas, la segunda urna contiene 2 bolitas negras y 2 bolitas blancas y la tercera contiene 3 bolitas negras y 1 bolita blanca. Nuestro experimento consiste en extraer una bolita de la primera urna y si sale blanca extraemos una segunda bolita de la tercera urna y si sale negra extraemos una segunda bolita de la segunda urna, ¿cuál es la probabilidad de obtener dos blancas?, ¿de obtener dos bolitas negras?, ¿de obtener una bolita de cada color?

El total de configuraciones posibles de dos bolitas, en este caso, es una combinación del principio de suma y de multiplicación. Si la primera bolita es blanca, hay 4 posibilidades para la segunda bolita y si la primera bolita es negra también, por lo que el total de posibilidades es 20. El total de configuraciones de dos bolitas blancas es $2 \cdot 1 = 2$ y el total de configuraciones de dos bolitas negras es $3 \cdot 2 = 6$ y, por lo tanto, el total de configuraciones de una bolita de cada color es 12. Es claro que las probabilidades coinciden con el experimento anterior. De hecho, queda claro que es otra manera de examinar el mismo experimento sólo que en vez de considerar la misma urna para la segunda extracción, tomamos otras urnas que contienen la misma cantidad de bolitas que debiera tener la urna al haber extraído una de ellas. Esto se aprecia gráficamente en la figura 4.13.

Parece claro, entonces, que todo lo que necesitamos para manipular las probabilidades en estos problemas de urnas y bolitas, son los conocimientos básicos de combinatoria y la regla de favorables sobre posibles.

Ahora construimos un experimento similar al anterior. Consideramos una vez más tres urnas, cada una de las cuales tiene en su interior 1 bolita negra y 1 bolita blanca. El experimento es como el que ya vimos, es decir, se extrae una bolita de la primera urna y si sale blanca extraemos una segunda bolita de la tercera urna y si sale negra extraemos una segunda bolita de la segunda urna.

Esto es equivalente a extraer 3 bolitas con restitución de una urna que tiene una bolita blanca y una negra. Por lo tanto, está claro que la probabilidad de obtener dos bolitas blancas es igual a la probabilidad de obtener dos bolitas negras y es igual a $\dfrac{1}{4}$ y, por lo tanto, la probabilidad de obtener una bolita de cada color es $\dfrac{1}{2}$ (ver figura 4.14 costado izquierdo).

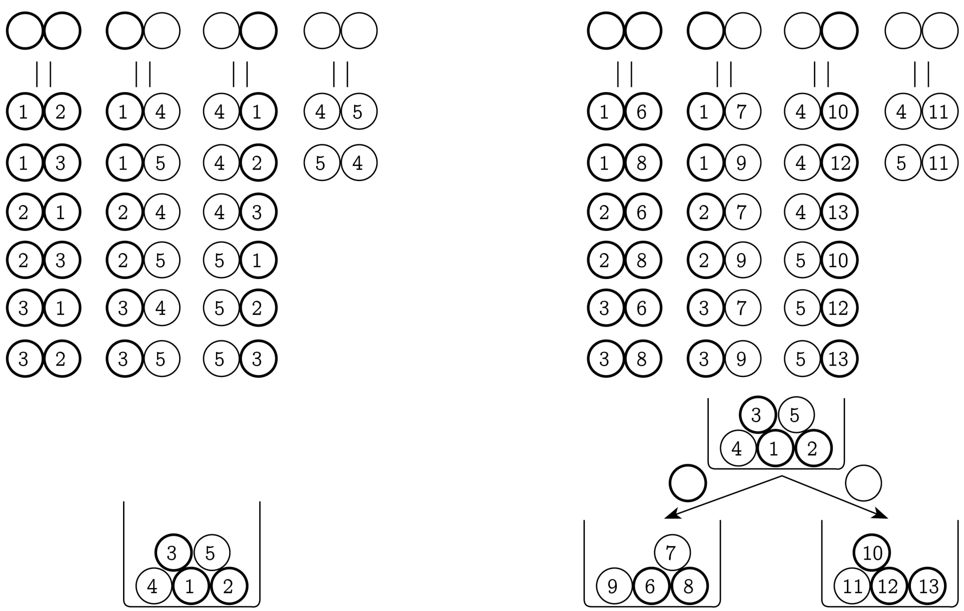


Figura 4.13. Experimentos equivalentes

Ahora bien, ¿cómo cambiarían las probabilidades si agregamos una bolita blanca y una negra a la tercera urna?. Es decir, si el total de bolitas en la tercera urna cambia pero la proporción entre bolitas blancas y negras sigue siendo la misma. La "intuición" nos dice que las probabilidades anteriores no debiesen cambiar.

Apliquemos una vez más la maquinaria que hemos estado utilizando hasta ahora. El total de configuraciones posibles es $1 \cdot 2 + 1 \cdot 4 = 6$. El total de favorables a obtener dos bolitas blancas es $1 \cdot 2$ y el total de favorables para obtener dos bolitas negras es $1 \cdot 1$ y, por lo tanto, el total de favorables para una de cada color es $6 - 2 - 1 = 3$. Pero entonces la probabilidad de obtener dos bolitas blancas es $\frac{2}{6} = \frac{1}{3}$, la probabilidad de obtener dos bolitas negras es $\frac{1}{6}$ y, ¡la probabilidad de obtener una bolita de cada color es $\frac{3}{6} = \frac{1}{2}$!.

Es decir ¿la probabilidad de obtener una bolita de cada color es la misma pero la probabilidad de obtener dos bolitas blancas aumentó y la probabilidad de obtener dos bolitas negras disminuyó?.

Entonces ¿nuestra intuición es incorrecta?, ¿o fueron nuestras cuentas las incorrectas?. Revisemos y precisemos nuestras cuentas un poco. Hay un total de 8 bolitas

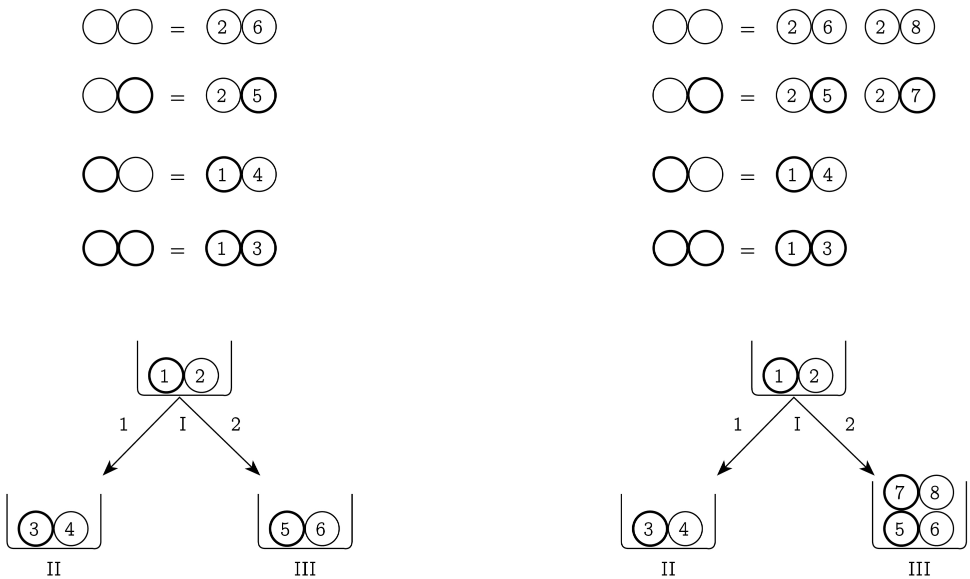


FIGURA 4.14. Paradoja de combinatoria y probabilidad

de las cuales 4 son blancas y 4 negras. Las vamos a numerar del 1 al 8 con las pares como bolitas blancas y las impares como bolitas negras, de modo que en la primera urna se encuentren las bolitas 1 y 2, en la segunda las bolitas 3 y 4 y en la tercera urna las bolitas 4,5,6,7 y 8.

Entonces, los resultados posibles de nuestro experimento son $1-3$,$1-4$,$2-5$,$2-6$,$2-7$,$2-8$ (ver figura 4.14 costado derecho). Al distinguir tan solo el color y codificando blanca como B y negro como N, obtenemos $N-N$, $N-B$, $B-N$, $B-B$, $B-N$ y $B-B$. Esto nos muestra que nuestras cuentas son correctas, ¿entonces es nuestro razonamiento el incorrecto?. La respuesta aunque parezca contradictoria es que nuestro razonamiento tampoco es incorrecto porque efectivamente las probabilidades son las mismas que existían antes de agregar las bolitas a la tercera urna, ¿cómo es esto posible?

Una vez más el culpable es la Falacia de Equiprobabilidad. En la primera urna, la probabilidad de obtener una bolita cualquiera es $\frac{1}{2}$, en la segunda urna, la probabilidad de obtener una bolita cualquiera es $\frac{1}{2}$, pero en la tercera urna la probabilidad de obtener una bolita cualquiera es $\frac{1}{4}$. Esto quiere decir, que la probabilidad de obtener $1-3$ o la probabilidad de obtener $1-4$, es en ambos casos, $\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$; pero la probabilidad de obtener $2-5$ es $\frac{1}{2}\cdot\frac{1}{4}=\frac{1}{8}$ al igual que la probabilidad de obtener $2-6$ ó $2-7$ ó $2-8$.

Es decir, *no todos los resultados son igualmente probables* por lo que la fórmula de favorables sobre posibles ya no tiene validez. Sin embargo, esto no quiere decir que no podamos calcular la probabilidad. La probabilidad de obtener dos bolitas blancas, que corresponde a obtener $2-6$ y $2-8$ al sumar ambas probabilidades nos da $2\cdot\frac{1}{8}=\frac{1}{4}$. La probabilidad de obtener dos bolitas negras, que corresponde a $1-3$ nos da $\frac{1}{4}$ y finalmente la probabilidad de obtener una bolita de cada color será $\frac{1}{2}$ lo que muestra que las probabilidades son las mismas que antes de agregar las bolitas a la tercera urna.

4.14.3 El juego de las nueve bolitas

A continuación, a través de un ejemplo concreto y simple, vamos a dar una pincelada al tipo de problemas y técnicas que hemos desarrollado y seguiremos desarrollando más adelante. La idea es considerar un ejemplo simple pero que contenga una cantidad suficiente de resultados para que las preguntas que hagamos no se puedan resolver de manera exhaustiva, al menos para un simple ser humano.

Consideremos una urna con nueve bolitas que estén rotuladas con los dígitos 1,2,3,4,5,6,7,8 y 9. Una extracción de bolitas de la urna es como seleccionar al azar un dígito entre 1 y 9. Nuestro experimento consistirá en extraer todas las bolitas de dicha urna sin restitución. Con las reglas del juego que hemos establecido esto se debe hacer una por una. Algunos de los resultados posibles se muestran en la figura 4.15.

Con nuestro experimento podemos poner a prueba los conocimientos de combinatoria para calcular probabilidades. El total de resultados posibles por el principio de multiplicación es $9!$ y por lo tanto cualquier resultado tiene probabilidad $\frac{1}{9!}$ de ocurrir.

Por ejemplo, veamos como calcular la probabilidad de que todas las bolitas rotuladas pares salgan antes que las bolitas rotuladas impares. Necesitamos el total de configuraciones favorables. Tenemos que considerar las secuencias de 9 dígitos que tienen en las primeras 4 posiciones un número par y en las 5 últimas un dígito impar. Es decir $4!\cdot 5!$ lo que nos da una probabilidad de $\frac{5!\cdot 4!}{9!}=\frac{1}{126}$ que es bastante pequeño.

En el espíritu del origen de las probabilidades, vamos a proponer un juego de azar derivado de este experimento. Se ofrece a los jugadores apostar a alguno de los siguientes 9 resultados. Dado $i\in\{1,\dots,9\}$ se dice que hay una *coincidencia en la extracción i* (que notaremos como A_i) si en la extracción i se obtiene la bolita rotulada con el dígito i.

Inicialmente nos vamos a concentrar en responder una pregunta que parece bastante simple, ¿cuál es la probabilidad de que haya alguna coincidencia?. Este evento,

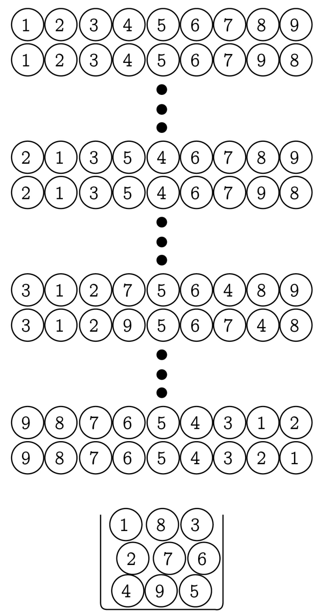


FIGURA 4.15. Extracción sin restitución de nueve bolitas

al estar descrito en palabras (en castellano) ya tiene una dificultad adicional que es propia al lenguaje y que produce confusiones al tratar de escribirla en términos matemáticos. Los conectivos lógicos "y" y "o" en castellano tienen significados ambiguos que causan estragos al ser llevados al contexto de las matemáticas. En este caso lo que estamos diciendo es cuáles son las secuencias en que para algún i se tiene coincidencia en la extracción i. Esto no quiere decir que debe ser sólo para un i.

Estudiemos inicialmente los eventos A_i. Por construcción asumimos que todas las secuencias son igualmente probables, pero no necesariamente los eventos A_i son todos equiprobables. Para calcular la probabilidad de A_i sólo necesitamos calcular las configuraciones favorables a A_i que son evidentemente $8!$ puesto que la única restricción es que en la extracción i se obtenga la bolita rotulada como i. Es decir, la probabilidad de que ocurra A_i es $\dfrac{1}{9}$ para cualquier valor $i \in \{1, \ldots, 9\}$. Esto quiere decir que da lo mismo apostar a cualquiera de estos eventos porque son todos equiprobables.

Un cálculo directo al sumar la probabilidad de cada uno de los A_i nos da $9 \cdot \dfrac{1}{9} = 1$. Entonces esto quiere decir que ¿la probabilidad de que ocurra alguno los A_i es 1? o lo que es lo mismo (en este caso), ¿siempre ocurre alguno de los eventos A_i?. Claramente no, puesto que la secuencia $2-3-4-5-6-7-8-1$ no es favorable a ninguno de los A_i.

El error cometido es simple (y muy común además) y proviene del hecho que se contó de mala manera. Por ejemplo, al contar las resultados favorables para A_1 y luego sumar los resultados favorables para A_2 no estamos contando los resultados favorables para A_1 o A_2 porque estamos contando más que eso. Lo que ocurre es que estamos contando algunos resultados dos veces, como por ejemplo, el resultado $1-2-3-4-5-6-7-8-9$ que es favorable a ambos resultados. Nos volvemos a encontrar con el mismo tipo de desafíos que resolvimos en la sección de combinatoria.

Intentemos contar una vez más los sucesos favorables para obtener A_1 o A_2. Con la misma idea de la combinatoria, podemos intentar descomponer estos resultados en otros disjuntos más fáciles de manejar. Es claro que los resultados favorables a A_1 o A_2 se pueden separar en tres grupos disjuntos. El grupo de resultados favorables a A_1 y no favorables para A_2, el grupo de resultados favorables a A_2 y no favorables para A_1 y finalmente el grupo de resultados favorables para A_1 y para A_2. La ventaja de esta descomposición es que estos tres grupos si los podemos contar fácilmente. En efecto, para ser favorable a A_1 y no a A_2 debo obtener la bolita rotulada 1 en la primera extracción y no la bolita rotulada 2 en la segunda extracción, es decir, $7 \cdot 7!$. El caso de resultados favorables para A_2 y no favorables para A_1 es análogo y nos da también $7 \cdot 7!$. El grupo de resultados favorables tanto para A_1 como para A_2 fija las dos primeras bolitas que deben ser extraídas y por lo tanto son $7!$. Entonces por el principio de suma, el número de resultados favorables para A_1 o A_2 es $(2 \cdot 7+1) \cdot 7!$ lo que implica que la probabilidad de que ocurra A_1 o A_2 es $\dfrac{(2 \cdot 7+1) \cdot 7!}{9!} = \dfrac{5}{24} < \dfrac{2}{9}$. El error cometido anteriormente, es decir, la suma de las favorables para A_1 y las favorables para A_2 menos las favorables para A_1 o A_2 nos da $2 \cdot 8! - (2 \cdot 7+1) \cdot 7! = 7!$ que coincide con las favorables para A_1 y A_2 (¿por qué?).

Podríamos intentar el mismo tipo de razonamiento para obtener el número de resultados favorables para A_1 o A_2 o A_3. Por desgracia, en ese caso, tendríamos que separar en más grupos; los resultados favorables sólo a uno de los tres eventos, los resultados favorables sólo a dos de los tres eventos y los resultados favorables a los tres eventos. Es decir $3+3+1=7$ grupos. Repitiendo el mismo tipo de argumentos esto nos da $3 \cdot 7 \cdot 6 \cdot 6! + 3 \cdot 6 \cdot 6! + 6!$ resultados favorables o sea un probabilidad de

$$\frac{3 \cdot 7 \cdot 6 \cdot 6! + 3 \cdot 6 \cdot 6! + 6!}{9!} = \frac{145}{504} < \frac{3}{9}.$$

Es cierto que es posible realizar lo mismo para los resultados favorables para A_1 ó A_2 ó A_3 ó A_4, pero ya se empieza a poner un poco escabroso. Al parecer el evento que queremos calcular es fácil de enunciar pero no tan fácil de escribir. En todo caso vemos que es posible hacerlo de este modo con un poco de paciencia y una buena computadora.

Intentemos otro enfoque inspirados en el problema de los prisioneros. El número de resultados favorables para el evento A_1 es $8!$ y por lo tanto el número de resultados no favorables para A_1 es $9! - 8!$. El número de resultados no favorables para A_1 pero si para A_2 es entonces $8! - 7!$ (como si consideramos una bolita menos en la urna).

Entonces el número de resultados no favorables ni para A_1 ni para A_2 es $9! - 2 \cdot 8! + 7!$. Como en el caso de los prisioneros obtenemos finalmente que el número de resultados favorables para ninguno de los eventos A_i es:

$$9! - \binom{9}{1}8! + \binom{9}{2}7! - \binom{9}{3}6! + \binom{9}{4}5! - \binom{9}{5}4! + \binom{9}{6}3! - \binom{9}{7}2! + \binom{9}{8}1!$$

Por lo tanto la probabilidad de que ninguno de los eventos ocurra es $\dfrac{3,641}{5,760}$ y por tanto la probabilidad de que ocurra alguno de los eventos A_i es $1 - \dfrac{3,641}{5,760} = \dfrac{2,119}{5,760}$. Esto quiere decir que lo más probable es que no ocurra ninguno de los eventos A_i.

Lo anterior muestra que con un ejemplo básico (1 urna y 9 bolitas) los cálculos asociados a preguntas simples pueden resultar bastante complejos. Es importante notar que los argumentos que utilizamos anteriormente, se aplican en general y que no son más que una aplicación directa de los principios de suma y multiplicación y la fórmula de favorables sobre posibles. La dificultad en este ejemplo es sólo de cálculo y la dificultad reside por ende exclusivamente en la aritmética involucrada.

Sin embargo, podemos hacer preguntas más complejas en este mismo experimento, como estudiar la causalidad entre eventos distintos. Por ejemplo, como cambian los cálculos si se apostó a A_2 y en la primera extracción se obtiene la bolita rotulada como 1 (es decir el resultado ya es favorable para A_1). ¿Es ahora más o menos probable que el resultado sea favorable para A_2? (pregunta crucial para un apostador si se puede retirar después de la primera extracción). Un razonamiento heurístico será decir que dado que ya se logró una vez una coincidencia, parece ser que va a ser más difícil volver a lograr otra coincidencia por lo que la probabilidad debiera ser menor. Veamos ahora si ese razonamiento tiene sentido matemático.

Esto no es tan difícil de calcular dado que el total de resultados posibles es ahora 8!, porque debe salir la bolita 1 en la primera extracción. De estos hay 7! favorables para A_2 por lo que la probabilidad es ahora $\dfrac{1}{8} > \dfrac{1}{9}$. Es decir, de entre los resultados favorables para A_1 la proporción de resultados favorables para A_1 y a A_2 es mayor que la proporción de resultados favorables para A_2 del total de resultados. Esto muestra que al contrario de lo que se razonó anteriormente, la probabilidad de obtener una coincidencia en la segunda extracción dado que se obtuvo una coincidencia en la primera, es mayor que antes.

Entonces el resultado A_1 influye favorablemente sobre el resultado A_2 con respecto a la probabilidad de ocurrencia. Entonces, ¿qué ocurre en caso contrario?. Lo único que sabemos es que la bolita 1 no salió en la primera extracción. Esto no nos dice exactamente que bolita es la obtenida en la primera extracción, sólo sabemos que no es la bolita 1. ¿Cuál es la probabilidad de que el resultado sea favorable para A_2 en este caso?.

El total de resultados posibles es $8 \cdot 8!$ puesto que no debe salir la bolita 1 en la primera extracción. Al sumar esto con los resultados posibles en el caso anterior obtenemos $8 \cdot 8! + 8! = 9!$ que es total de resultados posibles. Esto nos muestra que con probabilidad $\frac{1}{9}$ hay coincidencia en la primera extracción y con probabilidad $\frac{8}{9}$ no hay coincidencia en la primera extracción.

Nos conviene más ver los resultados que no son favorables para A_2. Hay $8!$ resultados que tienen a la bolita 2 en la primera extracción y además hay $7 \cdot 7 \cdot 7!$ que no tienen ni a la bolita 1 ni a la bolita 2 en la primera extracción y no tienen a la bolita 2 en la segunda extracción. Entonces, la probabilidad pedida es $1 - \frac{8! + 49 \cdot 7!}{8 \cdot 8!} = \frac{7}{64} < \frac{1}{9}$. Es decir ocurre lo contrario a la situación anterior ahora la probabilidad de obtener una coincidencia en la segunda extracción es menor que antes.

Entonces, el valor de la probabilidad de obtener una coincidencia en la segunda extracción, está entre el valor de la probabilidad de coincidencia cuando hay coincidencia en la primera extracción y el valor de la probabilidad de coincidencia cuando no hay coincidencia en la primera extracción. Para ser más precisos $\frac{1}{9} = \frac{1}{9} \cdot \frac{1}{8} + \frac{8}{9} \cdot \frac{7}{64}$. Es decir, que la probabilidad global es un *promedio ponderado de los dos casos y el ponderador es la probabilidad de que ocurran cada uno de esos dos casos.*

Podemos complicar un poco más las cosas y establecer una relación de causalidad inversa. Por ejemplo, ¿cuál es la probabilidad de que haya ocurrido A_1 sabiendo que ocurrió A_2?. Esto parece un poco confuso dado que A_1 se decide en la primera extracción y A_2 en la segunda extracción y por lo tanto A_1 ya se decidió si sabemos lo que pasó con A_2. En el fondo no es tan confuso si pensamos que cuando se realiza el juego, no miramos el resultado de la primera extracción cuando esta se realiza.

Originalmente la probabilidad de que ocurriese A_1 era $\frac{1}{9}$ pero ahora sabemos que ocurrió A_2 y la pregunta es: ¿ cambia eso la probabilidad de ocurrencia de A_1?. Vimos anteriormente que A_1 y A_2 tenían alguna influencia entre ellos por lo que la respuesta no parece ser que sigue siendo $\frac{1}{9}$. En efecto, el total de resultados favorables a A_1 es $8!$ y de estos resultados favorables a A_1 el total de resultados favorables a A_2 es $7!$ por lo que la probabilidad de que ocurra A_1 dado que ocurre A_2 es $\frac{1}{8}$!. Es decir, es más probable que haya una coincidencia en la primera extracción si hay una coincidencia en la segunda extracción.

Capítulo 5: Conceptos Básicos de Probabilidad

En este capítulo vamos a introducir el formalismo clásico de la Teoría de Probabilidades con Teoría de Conjuntos establecido por Kolgomorov. La idea es aprovechar el trabajo de los capítulos anteriores para hacer más intuitivas las definiciones y propiedades que vamos a introducir.

5.1 Experimento y Espacio Muestral

El fenómeno a modelar se llama usualmente *Experimento* y se denota como $\mathcal{E}$. Un Experimento es, normalmente, una acción que produce algún resultado. Para hacerlo que sea "aleatorio" se supone que, el experimento, debe arrojar más de un resultado. Normalmente el único control que se tiene sobre $\mathcal{E}$ es poder realizarlo.

Como ya se discutió anteriormente, los ejemplos clásicos corresponden a lanzar un dado de seis caras sobre una mesa, o, una moneda al aire. Sin disponer de más información ambos experimentos son considerados aleatorios, no por la naturaleza del experimento, sino por el hecho que no se sabe con exactitud cuál va a ser el resultado del Experimento. La fecha de cumpleaños de una persona cualquiera es un dato preciso, no aleatorio, pero es un evento aleatorio para alguien que no la conozca. Un mismo Experimento puede ser visto además de maneras distintas según las interrogantes que se tenga respecto a éste. Normalmente, al lanzar un dado, lo que interesa es saber el número que se obtiene en la cara superior, pero, nos podría interesar saber si el dado cae o no sobre la mesa. Entonces el mismo Experimento define otra situación aleatoria porque la pregunta que se quiere responder es otra.

La mejor representación o, más bien, la más adecuada, de los resultados posibles está en manos del que modela y se conoce como el *Espacio Muestral* Ω. El Espacio Muestral representa una codificación de los resultados posibles de $\mathcal{E}$ adecuada a lo que se quiere resolver.

Ejemplos:

1. $\mathcal{E}$ = "lanzar un dado", $\Omega = \{1,2,3,4,5,6\}$.
2. $\mathcal{E}$ = "lanzar una moneda", $\Omega = \{cara, sello\} = \{0, 1\}$.

Una vez que se fijó el Espacio Muestral una realización del experimento se simula como obtener un elemento ω que pertenece a Ω. Esto es una reducción drástica de información. En el caso del dado, con el espacio muestral del ejemplo anterior, no se sabe si cayó en la mesa o no, que forma tenía el dado si tenía algún color, etc...

En la Teoría de Probabilidades hay tres categorías de espacios muestrales que se estudian con técnicas diferentes.

1. **Finito:** El espacio muestral consiste de una cantidad finita de resultados distintos posibles.
2. **Numerable:** El espacio muestral consiste en un familia indexada por los naturales de resultados distintos posibles.
3. **No Numerable:** El espacio muestral consiste en una infinidad de resultados distintos posibles que no se pueden indexar por los naturales.

Nota: En este texto sólo trabajamos con Espacios Finitos que son los que tratamos en los capítulos previos, pero lo cierto es que la axiomática que introducimos en este capítulo no es más compleja en los otros casos.

El espacio muestral Ω es importante y no siempre se tiene una "receta" para calcularlo. Lo que define si el Espacio Muestral es adecuado es su capacidad para resolver el problema planteado.

5.1.1 Afirmaciones y Sucesos

Para escoger el Espacio Muestral no es suficiente tener definido $\mathcal{E}$, además es necesario tener claro las preguntas o afirmaciones que han motivado modelar dicho experimento. Normalmente las afirmaciones se notan con las letras α, β, etc... En el experimento $\mathcal{E}$ = "Lanzar un dado de seis lados" dos afirmaciones serían.

α="el número de la cara superior es 6" o bien β="el dado cayó en la mesa".

La afirmación α sugiere un espacio muestral $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$ que toma en cuenta todos los resultados posibles de la cara superior del dado. Por otro lado, la afirmación β sugiere el espacio muestral $\Omega_2 = \{0, 1\}$. Codificando como 1 si el dado cayó en la mesa y 0 cuando no cae en la mesa. En ambos casos hubo una drástica reducción de información respecto al experimento.

Intuitivamente el espacio muestral Ω_2 es "adecuado" para la afirmación β y el espacio muestral Ω_1 no lo es. Más formalmente, β es una afirmación válida para Ω_2 pero no para Ω_1. Una afirmación es válida para un experimento si, una vez que realizado, se puede dar un valor de verdad a dicha afirmación. En nuestro caso, saber el número obtenido en la cara superior no determina si el dado cayó en la mesa y, viceversa, saber si cayó o no en la mesa no determina el valor de la cara superior. Interesa, entonces, escoger un espacio muestral que haga válidas las respuestas a las preguntas que se quieran responder.

Nota: Esto corresponde a las preguntas que quedan determinadas en las regiones de un disco para el caso 1.

En el ejemplo anterior, cada una de las afirmaciones en el Espacio Muestral, en el que son válidas, tienen asociadas un subconjunto del Espacio Muestral. La afirmación α tiene asociado el subconjunto $A = \{6\}$ de Ω_1 y β tiene asociado el subconjunto $B = \{1\}$ de Ω_2. Esto porque una afirmación válida descompone ("particiona") al espacio muestral en dos conjuntos disjuntos. Un conjunto corresponde a los elementos del Espacio Muestral que hacen que la afirmación sea cierta y el segundo subconjunto está formado por los elementos que hacen que la afirmación sea falsa.

Por lo tanto a cada afirmación le corresponde un subconjunto de Ω formado por aquellos elementos de Ω que la hacen verdadera. A estos conjuntos les llamaremos *sucesos*. En consecuencia, todas las afirmaciones válidas tienen una correspondencia uno a uno con todos los subconjuntos posibles del espacio muestral que se conoce como las partes de Ω y se nota como $\mathcal{P}(\Omega)$.

Esto quiere decir que cualquier operación lógica entre afirmaciones tiene su operación equivalente entre conjuntos. Sólo por enumerar algunas; negación con complemento,"o" con unión, "y" con intersección, "o exclusivo" con diferencia simétrica (ver Figura 5.1.1):

AFIRMACION	SUCESO
α	A
β	B
$\sim \alpha$	A^c
$\alpha \vee \beta$	$A \cup B$
$\alpha \wedge \beta$	$A \cap B$
$\alpha \veebar \beta$	$A \Delta B$

Figura 5.1. Relación entre Afirmación y Suceso

En el lenguaje de conjuntos un elemento de A de $\mathcal{P}(\Omega)$ se dice que *ocurre* si el valor $\omega \in \Omega$ obtenido al realizar el experimento está en A.

Nota: Cada suceso es una unión de regiones del disco y que ocurra quiere decir que el resultado es una de las etiquetas.

5.1.2 Familias de Subconjuntos

Sea Ω un espacio muestral asociado a un cierto experimento $\mathcal{E}$. El conjunto $\mathcal{P}(\Omega)$ contiene todos los sucesos posibles asociados a afirmaciones válidas de un experimento. En general, mientras más sucesos distintos tengamos más trabajo habrá que hacer y, por lo tanto, es razonable tratar de considerar el mínimo posible que nos permita resolver el problema. Cuando el espacio muestral es finito (digamos N elementos) entonces el conjunto de todos los subconjuntos que se nota como $\mathcal{P}(\Omega)$ también es finito (más precisamente consta de 2^N elementos como ya vimos en el Capítulo de Combinatoria).

Para Espacios Muestrales en las otras categorías 2 y 3, el conjunto $\mathcal{P}(\Omega)$ resulta ser una familia demasiado exhaustiva de sucesos y, se hace necesario seleccionar un subconjunto $\mathcal{F} \subseteq \mathcal{P}(\Omega)$.

Nota: El subconjunto de sucesos más trivial de todos sería $\{\phi, \Omega\}$ que no tiene ningún interés práctico.

Obviamente, para poder responder a combinaciones lógicas de afirmaciones, la familia de subconjuntos $\mathcal{F}$ ha de tener un mínimo de propiedades algebraicas.Es decir, una familia $\mathcal{F}$ es un álgebra sobre Ω si:

1. $\Omega \in \mathcal{F}$
2. $\forall A \in \mathcal{F},\ A^c \in \mathcal{F}$
3. $\forall A, B \in \mathcal{F},\ A \cup B \in \mathcal{F}$.

La propiedad 1 es claramente necesaria dado que al menos debemos contener a todos los sucesos posibles. Una afirmación α válida particiona al espacio muestral en dos sucesos ($X = A \uplus A^c$). Por lo tanto, la familia $\mathcal{F}$ debe ser cerrada para la operación de complemento lo que justifica 2. Al considerar dos afirmaciones válidas distintas α y β particionamos el espacio muestral en cuatro sucesos $\Omega = (A \cap B) \uplus (A \cap B^c) \uplus (A^c \cap B) \uplus (A^c \cap B^c)$ (ver figura 5.2).

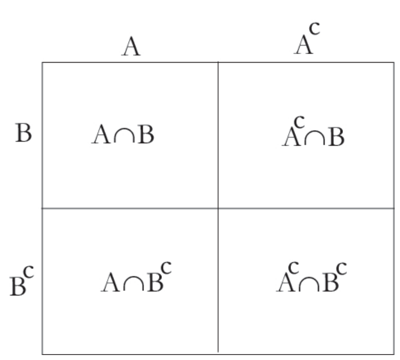


FIGURA 5.2. Partición por dos sucesos A y B

Nota: Esto corresponde a reetiquetar el Disco respecto a los resultados para los cuales la afirmación es válida o no y, para el caso de dos afirmaciones, los 4 resultados posibles y así sucesivamente para más resultados.

Es un ejercicio simple probar que basta considerar 2 y 3 para lograr que esos cuatro conjuntos estén en $\mathcal{F}$ cuando A y B están en $\mathcal{F}$. En realidad, cualquier conjunto que resulte de una cantidad finita de uniones e intersecciones, de elementos de $\mathcal{F}$ estará también en $\mathcal{F}$ cuando es un álgebra. En general no es difícil ver que la *intersección arbitraria* de álgebras es un álgebra.

En el caso en que $\Omega = \{1, \ldots, N\}$ el *álgebra mas pequeña* (en el sentido de la inclusión) que contiene a los sucesos $\{1\}$, $\{2\}$,...,$\{N\}$ resulta ser igual a la más grande que es $\mathcal{P}(\Omega)$. Es por eso que en los espacios muestrales en la categoría 1 se considera siempre $\mathcal{F} = \mathcal{P}(\Omega)$.

Si consideramos el experimento $\mathcal{E} =$ "Lanzar un dado $\mathbb{N}$ veces". Si nos interesa saber el número de caras obtenidas podemos considerar como espacio muestral $\Omega_1 = \mathbb{N}$ que es obviamente numerable. Las afirmaciones que sugieren este espacio muestral son $\alpha_n =$ "Salieron n caras" con $n \in \mathbb{N}$. Los sucesos asociados serán entonces $A_n = \{n\}$ con $n \in \mathbb{N}$.

Consideremos la afirmación $\alpha =$ "El número de caras obtenidas es par" que está asociada al conjunto $A = \{2, 4, \ldots, 2N, 2N+2, \ldots\}$. Este conjunto no se puede construir con un número finito de operaciones como unión e intersección de la familia de sucesos $\{A_1, \ldots, A_n, A_{n+1}, \ldots\}$. En efecto, es claro que $A = \biguplus_{n\in\mathbb{N}} A_{2n}$ y que $\biguplus_{n=0}^{N} A_{2n} \subset A$ para cualquier $N \in \mathbb{N}$. En este caso la más pequeña de las álgebras que contiene a la familia $\{A_n\}_{n\in\mathbb{N}}$ no contiene a A y por ende no es $\mathcal{P}(\Omega)$. Esto muestra que fuera de la categoría 1 es necesario manejar más que una cantidad finita de operaciones de unión e intersección. En estos casos le pedimos a $\mathcal{F}$ que sea una σ-álgebra de Ω que equivale a reemplazar la propiedad 3 por:

3σ. $\forall \{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}, \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$.

Al considerar una σ-álgebra cualquier conjunto que resulte de una cantidad numerable de uniones e intersecciones de elementos de $\mathcal{F}$ estará también en $\mathcal{F}$. La más pequeña de las σ-álgebras que contiene a la familia $\{A_n\}_{n\in\mathbb{N}}$ resulta ser la más grande es decir $\mathcal{P}(\Omega)$.

Volviendo al mismo experimento, si nos interesa saber si salió cara o sello en alguno de los lanzamientos, el Espacio Muestral considerado anteriormente no es adecuado. Uno sería $\Omega = \{0, 1\}^{\mathbb{N}}$ en el que el resultado $\{w_n\}_{n\in\mathbb{N}} \in \Omega$ se construye como $w_n = 1$ si en el n-ésimo lanzamiento se obtuvo cara y $w_n = 0$ si se obtuvo sello en el n-ésimo lanzamiento. Este espacio muestral no es numerable (de hecho tiene el mismo cardinal de los números reales). Las afirmaciones que dieron origen este Espacio Muestral son $\alpha_k =$ "Salió cara en el k-ésimo lanzamiento" con $k \in \mathbb{N}$ que tiene los sucesos asociados $A_k = \{\{w_n\}_{n\in\mathbb{N}} \in \Omega | w_k = 1\}$. La más pequeña de las σ-álgebras que contiene a esta familia de sucesos resulta no ser $\mathcal{P}(\Omega)$. Las ventajas de considerar esta σ-álgebra en vez de $\mathcal{P}(\Omega)$ será evidente al definir las medidas de probabilidad.

Resumiendo, dado un fenómeno que será modelado como una situación aleatoria lo primero que se define es un *Espacio de Probabilidad* $(\Omega, \mathcal{F})$ con Ω un conjunto y $\mathcal{F}$ una σ-álgebra de Ω. Cuando el cardinal de Ω es finito o numerable entonces $\mathcal{F} = \mathcal{P}(\Omega)$.

En general se tendrá una σ-álgebra $\mathcal{F}$ tal que $\{\phi, X\} \subseteq \mathcal{F} \subseteq \mathcal{P}(\Omega)$. Una vez que el espacio de probabilidad está claro es tiempo de introducir la medida de probabilidad.

5.2 Medida de Probabilidad

En este capítulo vamos a generalizar las nociones de probabilidad obtenidas para espacios infinitos. Es decir, la probabilidad de un suceso es un número entre 0 y 1 que mide cuán posible es que ocurra. A mayor valor de la probabilidad mayor es la posibilidad de que ocurra.

5.2.1 Caso Finito o Numerable

Cuando se tiene un espacio muestral finito $\Omega = \{1, \ldots, N\}$ y un suceso $A \in \mathcal{P}(\Omega)$. La máxima desinformación, y por ende lo mas aleatorio, es que cualquier $\omega \in \Omega$ sea igualmente probable. En este caso la probabilidad de que A ocurra será simplemente el cociente entre los casos favorables (los que pertenecen a A) y todos los casos posibles, es decir $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$. Esto se conoce como la *Fórmula de Laplace* en que el cálculo de las probabilidad se reduce a "contar" los sucesos favorables.

Nota: Esto corresponde a considerar un disco con todas las regiones con la misma área.

Consideremos el siguiente ejemplo.

Ejemplo 3. *Se tienen 4 bolitas en una urna, 2 bolitas negras y 2 bolitas rojas. Se toman dos bolitas de la urna ¿cuál es la probabilidad de que haya una bolita de cada color?.*

Una posible solución sería dar como espacio muestral $\Omega_1 = \{0, 1\}$ en que 0 representa no tener una bolita de cada color y 1 representa obtener una bolita de cada color. Entonces suponiendo que ambos sucesos son igualmente probables la probabilidad sería $\dfrac{1}{2}$.

Pero si consideramos $\Omega_2 = \{$"2 negras", "1 de cada color", "2 rojas"$\}$ entonces si los tres sucesos son igualmente probables, la probabilidad sería $\dfrac{1}{3}$.

Otra posibilidad sería considerar $\Omega_3 = \{0, 1\} \times \{0, 1\}$ en que 0 representa una bolita negra y 1 una bolita roja. En este caso los casos favorables son 2 de un total de 4 por lo que la probabilidad sería otra vez $\dfrac{1}{2}$.

Finalmente consideremos el espacio muestral

$$\Omega_4 = \{(a, b) \in \{1, 2, 3, 4\} \times \{1, 2, 3, 4\} | a \neq b\},$$

en que las bolitas negras son 1 y 2 y las rojas son 3 y 4. En este caso los casos favorables son 8 de un total de 12 por lo que la probabilidad sería en este caso $\frac{2}{3}$.

¿Cuál es la respuesta correcta? Estos problemas fueron los que generaron infinitas discusiones entre los matemáticos en el siglo XVII. El cálculo del cociente entre favorables y posibles, en las cuatro soluciones anteriores, es correcto y, podría decirse entonces, que todas las soluciones lo son. Sin embargo, el supuesto de que los singleton son equiprobables es compatible con la última solución. Esto porque el Espacio Muestral Ω_4 "simula" la realización del experimento dado que la primera coordenada representa la primera bolita extraída y la segunda coordenada representa la segunda bolita extraída. En estas condiciones es fácil verificar si el supuesto de equiprobable es o no razonable, sin tener que recurrir a algún razonamiento más complejo. En realidad, cualquiera de estos espacios muestrales podría modelar este experimento, pero, no suponiendo que los sucesos son igualmente probables. Para el primer espacio muestral la probabilidad de obtener un 0 debería ser $\frac{1}{3}$ y la de obtener un 1, $\frac{2}{3}$. En el segundo espacio muestral la probabilidad de obtener 2 negras es igual a la probabilidad de obtener 2 rojas e igual a $\frac{1}{6}$ y la probabilidad de obtener 1 de cada color $\frac{2}{3}$. Para el tercer espacio muestral la probabilidad de obtener $(0,0)$ es igual a la probabilidad de obtener $(1,1)$ que es igual a $\frac{1}{6}$ y la probabilidad de obtener $(0,1)$ es igual a la probabilidad de obtener $(1,0)$ que es igual a $\frac{2}{6}$. Es por lo tanto preferible tener un modelo probabilístico que nos permita modelar espacios muestrales no equiprobables.

Entonces, en general, dado un Experimento $\mathcal{E}$ y un Espacio de Probabilidad asociado $(\Omega, \mathcal{F})$, una manera de determinar cuáles son las posibilidades de que un suceso cualquiera de $\mathcal{F}$ ocurra es asociar, a cada elemento de $\mathcal{F}$, un número en $[0,1]$ que será mayor mientras más posibilidades tenga el suceso. Es decir, definir una función, que llamaremos $\mathbb{P}$ desde $\mathcal{F}$ al intervalo $[0,1]$. Para que los resultados sean consistentes esta función deberá tener algunas propiedades simples que la definen como una *medida de probabilidad*.

Una medida de probabilidad sobre $(\Omega, \mathcal{F})$ es una función $\mathbb{P} : \mathcal{F} \to [0,1]$ que satisface:

1. $\mathbb{P}(\Omega) = 1$.
2. $\forall \{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}, A_n \cap A_m = \phi \, (\forall n \neq m) : \quad \mathbb{P}\left(\biguplus_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$.

La suma que aparece en 2 se interpreta como:

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \sup_{N\in\mathbb{N}} \sum_{n=1}^{N} \mathbb{P}(A_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mathbb{P}(A_n).$$

Observación: Es importante notar que la función $\mathbb{P}$ está definida sobre $\mathcal{F}$ y <u>NO sobre Ω</u>.

Esto formaliza la noción intuitiva que tenemos de lo que es la probabilidad.

Nota: La propiedad (1) corresponde a que el área total del Disco es 1. La propiedad (2) corresponde a que el área de la unión es la suma de las áreas que la forman.

La construcción de un modelo de probabilidad es entonces asociar a un experimento $\mathcal{E}$ con ciertas afirmaciones $\alpha, \beta, \gamma, ...$ un espacio de probabilidad y, una medida de probabilidad sobre ese espacio. Es decir, $(\Omega, \mathcal{F}, \mathbb{P})$.

Este proceso nos permite modelar, es decir, codificar y obtener sólo la información relevante del experimento relativa a las afirmaciones que se quieren validar. Esto normalmente conlleva una reducción drástica de información y, en algunos casos, asumir ciertos supuestos que darán una solución aproximada al problema planteado. Esto es, sólo un ejemplo más de una modelación matemática, de un fenómeno con herramientas probabilistas.

5.2.2 Espacios Muestrales Finitos

La manera estándar para construir una medida de probabilidad, es definirla para una familia de sucesos $\mathcal{C} \subseteq \mathcal{F}$ de tal manera que se pueda extender de una única manera para el resto de los sucesos de $\mathcal{F}$. Normalmente se busca una familia de sucesos $\mathcal{C}$ para la cual cualquier elemento de $\mathcal{F}$ se escriba como una unión disjunta de elementos de $\mathcal{C}$. Al definir el valor de la probabilidad de estos sucesos a través de la propiedad 2 se extenderán de manera consistente a todos los elementos de $\mathcal{F}$.

Nota: Esto corresponde a definir el área de las regiones del Disco.

Si el espacio muestral Ω es finito o numerable veremos que basta definir la función de probabilidad sobre $\mathcal{C} = \{\{w\} \in \mathcal{P}(\Omega) | \omega \in \Omega\}$. Es decir dos medidas de probabilidad sobre un mismo espacio de probabilidad que coinciden sobre $\mathcal{C}$ son la misma medida de probabilidad. Formalmente:

Teorema 5.1. *Dado $(\Omega, \mathcal{P}(\Omega))$ finito o numerable y dos funciones de probabilidad $\mathbb{P}_1$, $\mathbb{P}_2$ tales que para cada ω en Ω, $\mathbb{P}_1(\{\omega\}) = \mathbb{P}_2(\{\omega\})$ entonces $\mathbb{P}_1 = \mathbb{P}_2$.*

Demostración. Cualquier A en $\mathcal{P}(\Omega)$ se escribe como $A = \biguplus_{\omega \in A} \{\omega\}$ con una unión a lo más numerable, por lo tanto por la propiedad 2 se tendrá que:

$$\mathbb{P}_1(A) = \mathbb{P}_1 \left(\biguplus_{\omega \in A} \{\omega\} \right) = \sum_{\omega \in A} \mathbb{P}_1(\{\omega\}) = \sum_{\omega \in A} \mathbb{P}_2(\{\omega\}) = \mathbb{P}_2 \left(\biguplus_{\omega \in A} \{\omega\} \right) = \mathbb{P}_2(A).$$

$\square$

Ahora veamos como deben definirse los valores sobre los singleton para obtener una medida de probabilidad, al extenderla sobre todas las partes. Las condiciones se resumen en el siguiente teorema:

Teorema 5.2. *Si Ω es finito o numerable una función $\mathbb{P}$ definida sobre $\{\{\omega\}/\omega \in \Omega\}$ define una única medida de probabilidad sobre $\mathcal{P}(\Omega)$ al extenderse como $\mathbb{P}(A) = \sum\limits_{\omega \in A} \mathbb{P}(\{\omega\})$ para cada A en $\mathcal{P}(A)$ sí y solo sí:*

$$\forall \omega \in \Omega \quad \mathbb{P}(\{\omega\}) \geq 0, \quad \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 1.$$

Demostración. La unicidad ya se demostró (ver Teorema 5.1). Por definición $\mathbb{P}(\Omega) = \sum\limits_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = 1$.

Dado $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{P}(\Omega)$ familia disjunta dos a dos, entonces:

$$\mathbb{P}\left(\biguplus_{n \in \mathbb{N}} A_n\right) = \sum_{\omega \in \biguplus\limits_{n \in \mathbb{N}} A_n} \mathbb{P}(\{\omega\}) = \sum_{n \in \mathbb{N}} \sum_{\omega \in A_n} \mathbb{P}(\{\omega\}) = \sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$$

Por definición se tiene que A en $\mathcal{P}(\Omega)$, $0 \leq \mathbb{P}(A) \leq 1$. Lo que muestra que es una medida de probabilidad. $\square$

En general siempre podemos considerar (basta enumerar) $\Omega = \{1, \ldots, N\}$ con algún N en $\mathbb{N}$ con lo cual la función de probabilidad queda definida por N números $\{p_n\}_{n=1}^{N} \subseteq [0,1]$ tales que $\sum\limits_{n=1}^{N} p_n = 1$.

En un contexto de medida de probabilidad podemos definir una *medida equiprobable* como el análogo combinatorio de todos se cuenta por igual. Es decir:

Definición 1. *Dado $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ con Ω finito, se dice* equiprobable *si:*

$$\forall\, \omega, \omega' \in \Omega, \quad \mathbb{P}(\{\omega\} = \mathbb{P}(\{\omega'\}).$$

Esto implica que para cualquier A en $\mathcal{P}(\Omega)$, $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$.

Basta ver que $1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\biguplus\limits_{\omega \in \Omega} \omega\right) = \sum\limits_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = \mathbb{P}(\{\omega\})|\Omega|$ por lo que $\mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}$.

Para cualquier A, se tiene que $\mathbb{P}(A) = \mathbb{P}\left(\biguplus\limits_{\omega \in A} \omega\right) = \sum\limits_{\omega \in A} \mathbb{P}(\{\omega\}) = \frac{|A|}{|\Omega|}$.

Por ejemplo, si consideramos el lanzamiento de una moneda como nuestro experimento y nos interesa saber si salió cara o sello el espacio de probabilidad más simple será $(\{0,1\}, \mathcal{P}(\{0,1\}))$. Existe una familia de medidas de probabilidad que le podemos asociar a este espacio. En este ejemplo basta definir el valor de $\mathbb{P}(\{1\}) = p \in [0,1]$

para que la medida de probabilidad quede determinada en todos lados. Esto nos da una "infinidad" de posibilidades para modelar el lanzamiento de una moneda que se diferencian en la probabilidad de obtener una cara. Algunas posibilidades son las siguientes:

	$\mathbb{P}(\phi)$	$\mathbb{P}(\{0\})$	$\mathbb{P}(\{1\})$	$\mathbb{P}(\{0,1\})$
$p=0$	0	0	1	1
$p=\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$	1
$p=\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	1
$p=\frac{2}{3}$	0	$\frac{2}{3}$	$\frac{1}{3}$	1
$p=1$	0	1	0	1

El caso $p=0$ y $p=1$ indican que la probabilidad de obtener cara es 0 y 1 respectivamente. Es importante notar que esto no quiere decir que no es posible obtener cara cuando $p=0$ o sello cuando $p=1$, sino que la probabilidad de que esto ocurra es nula. Volveremos sobre este tema más adelante. El caso $p=\dfrac{1}{2}$ indica que no se tiene ninguna información sobre la moneda y por lo tanto se asume que es tan probable obtener cara como sello. En el caso $p=\dfrac{1}{3}$ se asume que el sello tiene el doble de posibilidades que la cara y $p=\dfrac{2}{3}$ es el caso en que la cara es dos veces más probable que el sello.

Un modelo probabilístico para el experimento de lanzar la moneda con las afirmaciones planteadas, se convierte simplemente en fijar un valor para p en el intervalo $[0,1]$. La pregunta sobre cual es el "mejor" valor para p dada una moneda "real" no nos concierne por el momento. Para resolver esta pregunta normalmente se utiliza la evidencia "empírica" del experimento. Es decir, repetir muchas veces el experimento de lanzar la moneda y, en función de los resultados obtenidos, obtener un valor "empírico" para p. Explicar esta definición empírica de probabilidad, muy pronto, tiende a generar una gran confusión introduciendo dos nociones de probabilidad. Una probabilidad "real" y una "empírica". Para evitar estas confusiones de definición y, con el fin de poder formalizar, correctamente, lo que se quiere decir con probabilidad empírica, es que este tipo de preguntas las dejaremos para más adelante.

Una de las grandes limitaciones de los modelos de probabilidad es que, en el caso en que Ω sea numerable no es posible definir una medida equiprobable. Esto quiere decir que el experimento de escoger un número natural al "azar" (equiprobable) no es posible de modelar con la teoría de probabilidades. Son necesarias algunas generalizaciones de esta teoría para lograrlo (Teoría de la Medida).

En efecto consideremos que $\Omega=\mathbb{N}$ y sea $c\in]0,1]$ tal que $\mathbb{P}(\{n\})=c$ para cualquier $n\in\mathbb{N}$.

Pero, por la propiedad (1) de $\mathbb{P}$, se tiene que $1=\mathbb{P}(\mathbb{N})=\sum\limits_{n\in\mathbb{N}}\mathbb{P}(\{n\})=\sum\limits_{n\in\mathbb{N}}c=\infty$ lo que es una contradicción.

5.3 Propiedades Básicas

Volvamos ahora a $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad cualquiera (no necesariamente con Ω finito). Veamos algunas propiedades básicas que satisface cualquier medida de probabilidad y que son consecuencia directa de la definición de esta.

1. $\forall\, A \in \mathcal{F}$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Es claro que $\Omega = A \uplus A^c$ por las propiedades 1 y 2 se tendrá que

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \uplus A^c) = \mathbb{P}(A) + \mathbb{P}(A^c).$$

Una de las consecuencias de esto es que $\mathbb{P}(\phi) = 0$. Es importante notar que $\mathbb{P}(A) = 0$ NO implica que $A = \phi$.

Nota: En Discos esto corresponde a reetiquetar respecto a si se obtiene el evento A o no. El área total queda separada en dos regiones cuyas áreas deben sumar 1.

Uno de los usos de esta fórmula surge cuando un suceso es complejo de descomponer y su complemento no lo es. Por ejemplo, se tienen 2000 productos y se quiere calcular la probabilidad de que haya al menos uno fallado. Suponiendo que tenemos la probabilidad de que cada uno esté fallado, describir el suceso anterior, puede resultar extenso. El complemento, sin embargo, ninguno fallado, es muy simple (la intersección de los complementos).

5.3.1 El Problema del Cumpleaños

Veamos como aplicar esto a un problema planteado en 1939 por un matemático llamado Richard von Mises, que se conoce, hoy en día, como el problema de los cumpleaños y que tiene una solución que es por demás sorprendente.

Ejemplo 4. *(Problema del Cumpleaños)*
¿Cuantas personas debe haber en una sala para que la probabilidad de que haya al menos dos de ellos de cumpleaños el mismo día sea mayor al $50\,\%$*?.*

Respuesta: Podemos "simplificar" el problema al considerar que los días de un año siempre son 365 (la respuesta es muy parecida en el caso de años con 366 días). Si hay N personas en la sala un espacio muestral adecuado será $\Omega = \{1, \ldots, 365\}^N$ en que cada coordenada representa el día de nacimiento de una persona en la sala.

El suceso $A =$"ningún par de personas está de cumpleaños el mismo día " no es fácil de escribir pero $A^c =$"todos están de cumpleaños en días distintos" es fácil de contar. Volviendo a la combinatoria, si pensamos que los días del año son casillas y asignar la fecha de cumpleaños a una persona es lanzar una bolita a una casilla en que aceptamos más de una bolita por casilla, tendremos que el total de configuraciones en que todas las personas están de cumpleaños en días distintos es $\dfrac{365!}{(365-N)!}$ de un total de 365^N. Por lo tanto $\mathbb{P}(A^c) = \dfrac{365!}{(365-N)!365^N}$ y por la propiedad 1, $\mathbb{P}(A) = 1 - \dfrac{365!}{(365-N)!365^N}$.

Al evaluar esta fórmula para $N = 20$ obtenemos ya que $\mathbb{P}(A) > \frac{1}{2}$. De hecho para $N \geq 60$ se tiene que la probabilidad es mayor al 99 %.

5.3.2 Propiedades de la Medida de Probabilidad

2. $\forall\, A, B \in \mathcal{F},\ A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.

Dado que $B = (B \setminus A) \uplus A$ por la propiedad 2

$$\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A) \geq \mathbb{P}(A).$$

Esta propiedad muestra que las medidas de probabilidad son monótonas al considerar la inclusión de conjuntos.

3. $\forall A, B \in \mathcal{F}\ \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Como $A \cup B = (B \setminus A) \uplus A$ y $B = (A \cap B) \uplus (B \setminus A)$ por la propiedad 2

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A),$$
$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A).$$

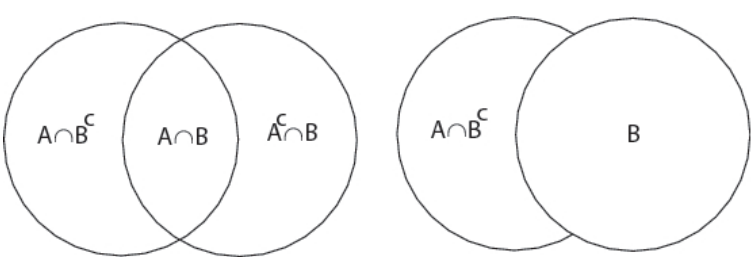


Figura 5.3. Diagrama de Venn

Despejando $\mathbb{P}(B \setminus A)$ de la segunda ecuación y reemplazándolo en la primera ecuación se obtiene el resultado. En el caso de 2 sucesos se puede comprender mejor utilizando un diagrama de Venn (ver figura 5.3).

¿Qué ocurre con la unión de tres sucesos?. Si agregamos $C \in \mathcal{F}$ razonando recursivamente obtenemos:

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C).$$

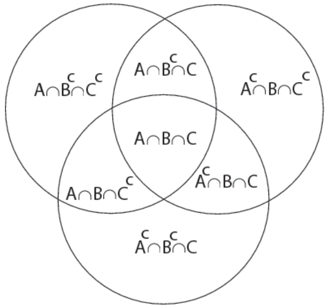


FIGURA 5.4. Diagrama de Venn

Esto también se puede representar gráficamente a través de un diagrama de Venn (ver figura 5.4). Por el mismo tipo de razonamiento recursivo la fórmula con cuatro elementos será entonces:

$$\begin{aligned}\mathbb{P}(A \cup B \cup C \cup D) &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) + \mathbb{P}(D) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) \\ &\quad - \mathbb{P}(A \cap D) - \mathbb{P}(B \cap C) - \mathbb{P}(B \cap D) - \mathbb{P}(C \cap D) + \\ &\quad + \mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap D) + \mathbb{P}(B \cap C \cap D) + \\ &\quad + \mathbb{P}(C \cap D \cap A) - \mathbb{P}(A \cap B \cap C \cap D).\end{aligned}$$

Esto no es posible representarlo de manera correcta a través de un diagrama de Venn (al menos en dos dimensiones).

4. Entonces si $A_1, \ldots, A_N \in \mathcal{F}$ entonces

$$\mathbb{P}(\bigcup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mathbb{P}(A_n) - \sum_{n>m} \mathbb{P}(A_n \cap A_m) + \cdots + (-1)^{N-1} \mathbb{P}(\bigcap_{n=1}^{N} A_n).$$

   Esto se conoce como fórmula de inclusión-exclusión que es del tipo que vimos en la sección de combinatoria.

5. $\forall\, A_1, \ldots, A_N \in \mathcal{F}$

$$\mathbb{P}(\bigcup_{n=1}^{N} A_n) \leq \sum_{n=1}^{N} \mathbb{P}(A_n).$$

   Si definimos $B_1 := A_1$ y $B_n := A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n$ para cada n en $\{1, \ldots, N\}$. Entonces:

$$\bigcup_{n=1}^{N} A_n = \biguplus_{n=1}^{N} B_n \Rightarrow \mathbb{P}\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mathbb{P}(B_n).$$

Como $B_n \subseteq A_n$ por la propiedad 5.3.2 $\mathbb{P}(B_n) \leq \mathbb{P}(A_n)$ y por lo tanto

$$\mathbb{P}\left(\bigcup_{n=1}^{N} A_n\right) \leq \sum_{n=1}^{N} \mathbb{P}(A_n).$$

El resultado también se puede demostrar por inducción sobre N (Ejercicio).

Como se puede apreciar en la demostración de todos los resultados anteriores la idea general, para calcular la probabilidad de un suceso complejo, es descomponerlo como unión disjunta de sucesos más simples. Es decir armar una *partición* de sucesos complejos en sucesos simples.

5.4 Probabilidad Condicional

Antes de proseguir con el desarrollo de la teoría de las probabilidades veamos, como motivación de esta sección, una pequeña historia que se conoce como el *dilema de los tres prisioneros.*

Ejemplo 5. *Un prisionero recibe la información de que uno de los prisioneros en su celda será ejecutado mañana. Con él son tres en la celda por lo que el prisionero considera que su probabilidad de morir es* $\frac{1}{3}$. *Por curiosidad le pregunta a uno de los guardias el nombre de alguno de los otros prisioneros que no será ejecutado mañana. Como el guardia sabe que eso no le da ninguna información al prisionero accede a darle el nombre. ¡Pero ahora el prisionero razona que su probabilidad de morir subió a* $\frac{1}{2}$*! y lamenta profundamente haber hecho la pregunta al guardia.*

Todos los cálculos del prisionero parecen correctos y sin embargo, ¿llevan a una contradicción?. En realidad no hay contradicción alguna, $\frac{1}{2}$ es la probabilidad de que el prisionero muera sabiendo que uno de los otros no muere que no es lo mismo que la probabilidad de que muera (que seguirá siendo $\frac{1}{3}$). Para realmente convencerse de que no hay tal contradicción es necesario hacer un poco más de teoría.

Una vez más consideremos un Espacio de Probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ cualquiera. Veremos, a continuación, cómo combinar distintos espacios de probabilidad e incorporar nueva información de manera de aprovechar el trabajo ya realizado. Todas las fórmulas y definiciones que se obtengan son sólo consecuencias de la definición de espacio de probabilidad. Primero veremos como se incorpora nueva información (no contradictoria con el modelo), con respecto al experimento, sin tener que volver a redefinir todo.

Supongamos que se tiene nueva información sobre el experimento, después de construido el modelo de probabilidad. Es posible que, esta nueva información, descarte el modelo construido por completo. Una ganancia de información compatible con el modelo sería por ejemplo que un cierto suceso B en $\mathcal{F}$ ocurrirá en la próxima realización del experimento o bien que la probabilidad de que B ocurra es 1 (que NO es lo mismo). ¿Cómo modifica esta información las probabilidades ?

Por ejemplo, en el capítulo de combinatoria se descomponía un procedimiento en varios procedimientos intermedios como técnica para poder contar. Esto equivale a fijar lo que sucederá, en alguna etapa intermedia, con alguno de los valores posibles y, dado eso, contar, para luego sumar sobre todos los posibles valores intermedios. De un punto de vista teórico no es necesario modificar el espacio de probabilidad, es decir, se conserva $(\Omega, \mathcal{F})$ (Lo "peor" que puede pasar es que se reduzcan los resultados posibles). Es claro, por otro lado, que la medida de probabilidad debe cambiar, al menos ahora la probabilidad que tiene B de ocurrir es 1.

La pregunta es ¿Cómo definir la nueva función de probabilidad, que notaremos $\mathbb{P}_B$ o bien $\mathbb{P}(\cdot|B)$?. Las posibilidades son muchas, por ejemplo, definir la probabilidad de B como 1 y definirla como 0 para cualquier otro suceso. Sin embargo, si se quiere incorporar la información que tenía el planteamiento anterior, las posibilidades se reducen. Veamos que es lo mínimo que se le debe pedir a esta nueva probabilidad. Es claro que para cualquier A en $\mathcal{F}$ tal que $\mathbb{P}(A \cap B) = 0$, $\mathbb{P}_B(A)$ deberá ser 0. En particular, para cualquier A con $A \cap B = \phi$. Es decir, los sucesos con probabilidad no nula deberán tener probabilidad no nula de intersectar a B. Formalmente esto equivale a exigir $\mathbb{P}_B(A) = \mathbb{P}_B(A \cap B)$. Dados C y D en $\mathcal{F}$ con $\mathbb{P}(C \cap B) \geq 0$ y $\mathbb{P}(D \cap B) \geq 0$. Nos gustaría mantener la proporción, es decir:

$$\frac{\mathbb{P}(C \cap B)}{\mathbb{P}(D \cap B)} = \frac{\mathbb{P}_B(C \cap B)}{\mathbb{P}_B(D \cap B)}.$$

Como además se debe tener que $\mathbb{P}_B(B) = 1$, al considerar $D = B$ se tiene que:

$$\frac{\mathbb{P}(C \cap B)}{\mathbb{P}(B)} = \mathbb{P}_B(C \cap B) = \mathbb{P}_B(C).$$

Lo anterior motiva la siguiente definición.

Dado B en $\mathcal{F}$ con $\mathbb{P}(B) \neq 0$ se define la *probabilidad condicional* de B como la medida de probabilidad, $\mathbb{P}_B : \mathcal{F} \to [0, 1]$ definida por:

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \forall A \in \mathcal{F}.$$

Nota: En discos esto corresponde a simplificar eliminando los resultados que no pertenecen a B.

Veamos que esto define una medida de probabilidad. Primero se tiene la propiedad 1 pues $\mathbb{P}_B(\Omega) = \frac{\mathbb{P}(\Omega\cap B)}{\mathbb{P}(B)} = 1$. Para probar la propiedad 2 sea $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ familia disjunta 2 a 2 entonces:

$$\begin{aligned}
\mathbb{P}_B\left(\biguplus_{n\in\mathbb{N}} A_n\right) &= \frac{\mathbb{P}(\left(\biguplus_{n\in\mathbb{N}} A_n\right)\cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\biguplus_{n\in\mathbb{N}} A_n \cap B)}{\mathbb{P}(B)} \\
&= \sum_{n\in\mathbb{N}} \frac{\mathbb{P}(A_n\cap B)}{\mathbb{P}(B)} = \sum_{n\in\mathbb{N}} \mathbb{P}_B(A_n).
\end{aligned}$$

Lo que demuestra que $\mathbb{P}_B$ es otra medida probabilidad sobre $(\Omega, \mathcal{F})$.

PREGUNTA: Es claro que $\mathbb{P}_\Omega = \mathbb{P}$. ¿Qué ocurre con $\mathbb{P}_B$ cuando $\mathbb{P}(B) = 1$ pero $B \neq \Omega$?.

Nota: La probabilidad condicional es una función de probabilidad con respecto a su primer argumento, no con respecto al segundo. En particular no es correcto $\mathbb{P}(A|B) + \mathbb{P}(A|B^c) = 1$.

Ahora veamos la influencia que tiene, que un suceso ocurra, sobre la probabilidad de otro. Dados $A, B \in \mathcal{F}$ con $\mathbb{P}(B) > 0$ decimos que B *influye* $\mathbb{P}$*-favorablemente sobre* A si $\mathbb{P}(A|B) \geq \mathbb{P}(A)$ y B *influye* $\mathbb{P}$*-desfavorablemente sobre* A si $\mathbb{P}(A|B) \leq \mathbb{P}(A)$. Si B influye $\mathbb{P}$-favorablemente y $\mathbb{P}$-desfavorablemente sobre A diremos que B *no influye por* $\mathbb{P}$ sobre A.

Si consideramos la relación influir $\mathbb{P}$-favorablemente (o $\mathbb{P}$-desfavorablemente o no influir por $\mathbb{P}$) sobre sucesos de probabilidad positiva entonces tenemos una relación de equivalencia. Es decir, refleja, simétrica y transitiva.

Veamos un ejemplo:

Ejemplo 6. *Se dispone de dos urnas con 3 bolas cada una. La primera, tiene 2 bolitas negras y 1 bola roja; La segunda tiene 1 bola negra y 2 bolitas rojas. Consideremos el experimento $\mathcal{E}$ = "Tomar una urna y luego, una bola de la urna elegida".*

Definamos los sucesos $A =$"sacar una bola negra" y B="tomé la primera urna ".

¿Cuál es la probabilidad de A? Sabiendo que A ocurrió, ¿cuál es la probabilidad de B?.

Consideremos $\Omega = \{1, 2\} \times \{0, 1\}$ donde $\{1, 2\}$ representa las urnas y $\{0, 1\}$ representa obtener una bola negra o 1 bola roja respectivamente. Por la propiedad (2):

$$\mathbb{P}(A) = \mathbb{P}((A \cap B) \uplus (A \cap B^c)) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c). \tag{5.1}$$

Pero además por la manera que se realiza el experimento:

$$\mathbb{P}(A|B) = \frac{2}{3} \quad \mathbb{P}(B) = \tfrac{1}{2}$$
$$\mathbb{P}(A|B^c) = \frac{1}{3} \quad \mathbb{P}(B^c) = \tfrac{1}{2}$$

Por lo tanto $\mathbb{P}(A \cap B) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$ y $\mathbb{P}(A \cap B^c) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$. Al reemplazar en la ecuación 5.1 tenemos que $\mathbb{P}(A) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$.

Como $\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \cdot \mathbb{P}(B)}{\mathbb{P}(A)}$ se tendrá que $\mathbb{P}(B|A) = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$.

Como conclusión B influía $\mathbb{P}$-favorablemente sobre A y B^c influía $\mathbb{P}$-desfavorablemente sobre A.

La probabilidad condicional es una herramienta poderosa, para poder calcular la intersección, cuando se tiene un relación de causalidad, entre los sucesos involucrados, como en el ejemplo anterior.

Por ejemplo, una fórmula para calcular la intersección de dos sucesos A y B en $\mathcal{F}$ de probabilidad positiva será:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B)$$

Si la causalidad es inversa entonces podemos intercambiar los roles de A y B y obtenemos:

$$\mathbb{P}(A \cap B) = \mathbb{P}(B|A) \cdot \mathbb{P}(A)$$

Notemos que esto es una generalización del Principio de Multiplicación que vimos en la sección de combinatoria.

Es posible extender la fórmula anterior a una intersección de más de dos elementos.

Lema 5.3. Probabilidad Producto.
$\forall A_1, \ldots, A_N \in \mathcal{F}$, $\mathbb{P}(A_1 \cap \cdots \cap A_N) > 0$:

$$\mathbb{P}(A_1 \cap \cdots \cap A_N) = \mathbb{P}(A_N | A_1 \cap \cdots \cap A_{N-1}) \cdot \cdots \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_1) \tag{5.2}$$

Demostración. Por definición se tiene el caso $N = 2$, razonando por inducción:

$$\mathbb{P}((A_1 \cap A_2) \cdots \cap A_N \cap A_{N+1})$$
$$\|$$
$$\mathbb{P}(A_{N+1}|(A_1 \cap A_2) \cdots \cap A_N) \cdot \cdots \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdot \mathbb{P}(A_1 \cap A_2)$$
$$\|$$
$$\mathbb{P}(A_{N+1}|(A_1 \cap A_2) \cdots \cap A_N) \cdot \cdots \cdot \mathbb{P}(A_3|A_1 \cap A_2) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_1).$$

□

La probabilidad condicional nos permite agregar información al modelo probabilístico al "forzar" que un suceso ocurra. Esto nos permite tener una generalización del principio de suma que vimos en la sección de combinatoria.

Lema 5.4. Probabilidades Totales.
$\forall A_1, \dots, A_N \in \mathcal{F}$ disjuntos 2 a 2 con $\mathbb{P}\left(\biguplus_{n=1}^{N} A_n\right) = 1$ y $\mathbb{P}(A_n) > 0$ para cualquier $n \in \{1, \dots, N\}$. Entonces $\forall A \in \mathcal{F}$

$$\mathbb{P}(A) = \sum_{n=1}^{N} \mathbb{P}(A|A_n) \cdot \mathbb{P}(A_n). \tag{5.3}$$

Demostración. Primero veamos que $\mathbb{P}(A) = \mathbb{P}(A \cap \biguplus_{n=1}^{N} A_n)$. Es claro que $A = (A \cap \biguplus_{n=1}^{N} A_n) \uplus (A \cap \bigcap_{n=1}^{N} A_n^c)$. Dado que $A \cap \bigcap_{n=1}^{N} A_n \subseteq \bigcap_{n=1}^{N} A_n$ y por hipótesis $\mathbb{P}(\bigcap_{n=1}^{N} A_n) = 0$ por la propiedad 5.3.2 se concluye. Como $\mathbb{P}(A \cap A_n) = \mathbb{P}(A|A_n)\mathbb{P}(A_n)$ se tiene que:

$$\mathbb{P}(A) = \sum_{n=1}^{N} \mathbb{P}(A \cap A_n) = \sum_{n=1}^{N} \mathbb{P}(A|A_n)\mathbb{P}(A_n).$$

□

El planteamiento inverso será: dado un suceso A y, una $\mathbb{P}$-partición de Ω, ¿cuán probable es que haya ocurrido alguno de los sucesos sabiendo que A ocurrió.

Teorema 5.5. Bayes(o de las causas). *$\forall A_1, \dots, A_N \in \mathcal{F}$ disjuntos 2 a 2 tales que $\mathbb{P}\left(\biguplus_{n=1}^{N} A_n\right) = 1$ entonces $\forall A \in \mathcal{F}$ tal que $\mathbb{P}(A) > 0$*

$$\mathbb{P}(A_m|A) = \frac{\mathbb{P}(A|A_m) \cdot \mathbb{P}(A_m)}{\sum_{n=1}^{N} \mathbb{P}(A|A_n) \cdot \mathbb{P}(A_n)}. \tag{5.4}$$

Demostración. Directamente de la definición de probabilidad condicional

$$\mathbb{P}(A_m|A) = \frac{\mathbb{P}(A_m \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|A_m) \cdot \mathbb{P}(A_m)}{\mathbb{P}(A)}.$$

Al aplicar probabilidades totales 5.3 se concluye que:

$$\mathbb{P}(A_m|A) = \frac{\mathbb{P}(A|A_m) \cdot \mathbb{P}(A_m)}{\sum_{n=1}^{N} \mathbb{P}(A|A_n)\mathbb{P}(A_n)}.$$

□

Nota: En Discos estos corresponde a invertir la causalidad con la Tabla de Bayes.

Veamos algunos ejemplos:

Ejemplo 7. *Se lanza un dado de seis caras y dos personas diferentes dicen que el resultado es 6. En general, se sabe que la primera persona miente con probabilidad* $p \in]0,1[$ *y la segunda persona miente con probabilidad* $r \in]0,1[$*. ¿Cuál es la probabilidad de que el resultado sea 6?, ¿para qué valores de* p *y* q *esta probabilidad sube a* $\frac{1}{2}$*?.*

Un Espacio Muestral adecuado sería $\Omega = \{0,1\} \times \{0,1\} \times \{0,1\}$. La primera coordenada vale 1 si salió 6 y 0 si no, la segunda coordenada vale 1 si la primera persona dijo la verdad y 0 si mintió y lo mismo para la tercera coordenada con la segunda persona.

Sea A, el suceso asociado a $\alpha =$"salió 6", B el suceso asociado a $\beta =$"la primera persona dijo que salió 6" y D, el suceso asociado a $\gamma =$"la segunda persona dijo que salió 6".

A *priori* se tiene que $\mathbb{P}(A) = \frac{1}{6}$. Además se tiene que $\mathbb{P}(B \cap D|A) = (1-p)(1-r)$, $\mathbb{P}(B \cap D|A^c) = p \cdot r$ por 5.3:

$$\mathbb{P}(B \cap D) = \mathbb{P}(B \cap D|A)\mathbb{P}(A) + \mathbb{P}(B \cap D|A^c)\mathbb{P}(A^c) = \frac{(1-p)(1-r)}{6} + \frac{p \cdot r}{6}.$$

La pregunta es, cuál es la probabilidad a *posteriori*, es decir, $\mathbb{P}(A|B \cap D)$ que por 5.4 resulta ser:

$$\mathbb{P}(A|B \cap D) = \frac{\frac{(1-p)(1-r)}{6}}{\frac{(1-p)(1-r)}{6} + \frac{p \cdot r}{6}} = \frac{(1-p)(1-r)}{(1-p)(1-r) + p \cdot r}.$$

Para que $\mathbb{P}(A|B \cap D) = \frac{1}{2}$ es necesario que $\mathbb{P}(B \cap D|A)\mathbb{P}(A) = \mathbb{P}(B \cap D|A^c)\mathbb{P}(A^c)$. Por lo tanto $(1-p)(1-r) = p \cdot r$ o lo que es lo mismo $\frac{1-p}{p} = \frac{r}{1-r}$ que implica que $r = 1 - p$.

Esto muestra que, con dos personas que mienten la mitad de las veces, la probabilidad de que sea 6, dado que las dos dicen que es 6 es $\frac{1}{2}$.

5.5 Los problemas de las Reinas y las Tarjetas II

En esta sección retomamos algunos problemas que resolvimos con la metodología de Discos.

5.5.1 El Problema de las Reinas II

Volvamos a resolver el problema de las Reinas que desarrollamos en la sección 3.2. Las figuras a las que se hace referencia para la resolución de este problema se encuentran en dicha sección.

Ejemplo 8. *Se tiene un tablero de ajedrez y se colocan dos reinas en el tablero (en posiciones distintas), ¿cuál es la probabilidad que se amenacen?, ¿cuál es la probabilidad que la primera reina este en la Región 3 (ver figura 3.1), dado que las reinas se amenazan?*

Consideremos el espacio muestral $\Omega = \{1, \ldots, 8\}^2 \times \{1, \ldots, 8\}^2$ en que la primera coordenada indica la posición de la primera reina lanzada y la segunda coordenada indica la posición de la segunda reina lanzada.

Nota: Se podría considerar un espacio muestral más pequeño dado que los elementos (i, i) con $i \in \{1, \ldots, 8\}$ no son resultados posibles dado el experimento. En este caso preferimos conservarlos y asumir que tienen probabilidad 0 de ocurrir.

Por el principio de multiplicación el número total de combinaciones es $64 \cdot 63$. Sea A el suceso definido por la afirmación $\alpha =$"Las reinas están amenazadas". La expresión exacta de A con respecto al espacio muestral no será necesaria para calcular $\mathbb{P}(A)$. Para las que no juegan ajedrez recordemos que una reina amenaza todas las casillas que están en la misma vertical, horizontal o cualquiera de las dos diagonales. En ajedrez, el control de las cuatro casillas centrales es algo deseable, porque éstas cuatro casillas son las que más casillas amenazan (27). La idea es utilizar probabilidades totales con respecto a la $\mathbb{P}$-partición dada por cuatro regiones disjuntas donde la primera reina puede caer (ver figura 3.1).

La ventaja de estas regiones, es que la probabilidad condicional de que las dos reinas estén amenazadas es la misma si la primera reina está en cualquier parte de una misma región. Si denotamos por A_i la probabilidad de que la primera reina esté en la región $i \in \{1, 2, 3, 4\}$, un sencillo cálculo combinatorio (favorables sobre posibles) nos da:

$$\begin{array}{ll}
\mathbb{P}(A|A_1) = \dfrac{21}{63} & \mathbb{P}(A_1) = \dfrac{28}{64} \\
\mathbb{P}(A|A_2) = \dfrac{23}{63} & \mathbb{P}(A_2) = \dfrac{20}{64} \\
\mathbb{P}(A|A_3) = \dfrac{25}{63} & \mathbb{P}(A_3) = \dfrac{12}{64} \\
\mathbb{P}(A|A_4) = \dfrac{27}{63} & \mathbb{P}(A_4) = \dfrac{4}{64}
\end{array}$$

Por lo tanto, por el lema 5.3, la probabilidad de que las reinas estén amenazadas es:

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A|A_1)\mathbb{P}(A_1) + \mathbb{P}(A|A_2)\mathbb{P}(A_2) + \mathbb{P}(A|A_3)\mathbb{P}(A_3) + \mathbb{P}(A|A_4)\mathbb{P}(A_4) \\
&= \frac{28 \cdot 21}{64 \cdot 63} + \frac{23 \cdot 20}{64 \cdot 63} + \frac{25 \cdot 12}{64 \cdot 63} + \frac{27 \cdot 4}{64 \cdot 63}
\end{aligned}$$

La siguiente pregunta es calcular $\mathbb{P}(A_3|A)$:

$$\mathbb{P}(A_3|A) = \frac{\mathbb{P}(A|A_3)\mathbb{P}(A_3)}{\mathbb{P}(A)} = \frac{\dfrac{25}{63} \cdot \dfrac{12}{64}}{\dfrac{28 \cdot 21}{64 \cdot 63} + \dfrac{23 \cdot 20}{64 \cdot 63} + \dfrac{25 \cdot 12}{64 \cdot 63} + \dfrac{27 \cdot 4}{64 \cdot 63}}$$

5.5.2 El Problema de las Tarjetas

Resolveremos, otra vez, el problema de las tarjetas enunciado en la sección 3.5.

Ejemplo 9. *Se tienen tres tarjetas, una roja por ambos lados, una blanca por ambos lados y una blanca por un lado y roja por el otro. Se toma una tarjeta al azar y se coloca sobre la mesa. Sabiendo que la cara superior es roja, ¿cuál es la probabilidad que la otra cara sea roja?*

Sea $\Omega = \Omega_1 \times \Omega_2^2 = \{1, 2, 3\} \times \{R, B\}^2$.Este espacio muestral considera primero la tarjeta que se obtiene y luego el color de cada una de sus dos caras (primero el color de la cara superior). La selección al azar se traduce en $\mathbb{P}(\{1\} \times \Omega_2^2) = \mathbb{P}(\{2\} \times \Omega_2^2) = \mathbb{P}(\{3\} \times \Omega_2^2) = \frac{1}{3}$ (Cuando esta claro el contexto se omite $\times\Omega_2^2$).

Si la numeración de las tarjetas sigue el orden del enunciado, entonces, $\mathbb{P}(R \times R|1) = 1$, $\mathbb{P}(B \times B|2) = 1$ y $\mathbb{P}(R \times B|3) = \mathbb{P}(B \times R|3) = \frac{1}{2}$.

¿Esto define una sola función de probabilidad ? En efecto, como el espacio es finito basta ver que está bien definida sobre los singleton.

Los únicos sucesos con probabilidad no nula (suman 1) son:

$$\mathbb{P}(1 \times R \times R) = \mathbb{P}(R \times R|1)\mathbb{P}(1) = \frac{1}{3}$$
$$\mathbb{P}(2 \times B \times B) = \mathbb{P}(B \times B|2)\mathbb{P}(2) = \frac{1}{3}$$
$$\mathbb{P}(3 \times R \times B) = \mathbb{P}(R \times B|1)\mathbb{P}(3) = \frac{1}{6}$$
$$\mathbb{P}(3 \times B \times R) = \mathbb{P}(B \times R|1)\mathbb{P}(3) = \frac{1}{6}$$

La pregunta era cuanto valía $\mathbb{P}(\times R|R\times)$, por 5.4:

$$\mathbb{P}(R \times R|R\times) = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6} + 0 + 0} = \frac{2}{3}$$

La intuición en general para este problema sugiere que la probabilidad debiera ser $\frac{1}{2}$ por la simetría entre los colores. Curiosamente, como ya mencionamos, si este problema es planteado de manera equivalente con bolitas y casillas la confusión es mucho menor. En efecto, si consideramos tres casillas (una con 2 bolitas blancas una con 2 bolitas rojas y, una con 1 bola blanca y una roja) y, sacamos al azar, una bolita de alguna de las urnas y obtenemos una bola roja, es razonable pensar que es más probable que la otra bolita sea roja, porque es más probable que haya seleccionado la segunda urna.

5.6 Independencia

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ una noción importante es la *independencia por* $\mathbb{P}$ de dos sucesos, digamos A y B, de probabilidad positiva (ambos). Intuitivamente se diría que dos sucesos son independientes, si el hecho de que ocurra, o no, uno de ellos no influye la probabilidad de que el otro ocurra. Es decir A y B son independientes por $\mathbb{P}$ si B influye favorable y desfavorablemente por $\mathbb{P}$ sobre A. Formalmente esto quiere decir que $\mathbb{P}(A|B) = \mathbb{P}(A)$ lo que equivale a $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Es claro que la noción de independencia por $\mathbb{P}$ entonces, depende de la medida de probabilidad.

Nota: En el contexto de discos sólo definimos independencia de Eventos Aleatorios y no de sucesos.

En el lenguaje corriente se diría que las afirmaciones "Hoy llueve" y "Me saco el Loto en el próximo sorteo" son independientes. Estas afirmaciones definen sucesos cuando se define un espacio muestral y estos sucesos tienen sus probabilidades una vez que se tiene la medida de probabilidad. ¿Qué pasa si la persona sólo compra boletos de Loto en los días de lluvia? En este caso las afirmaciones no son independientes, porque saber la respuesta a una de ellas modifica la probabilidad de que la otra afirmación sea cierta. Será necesario entonces "borrar" el significado coloquial que tenemos de la palabra independencia. Otro elemento que lleva a confusión es el siguiente: Dos sucesos disjuntos son independientes dado que "no tienen nada que ver". Este razonamiento es errado, por el contrario al ser disjuntos si uno ocurre el otro no ocurre por lo tanto se influyen $\mathbb{P}$-desfavorablemente de manera radical.

En realidad la independencia es más una noción entre afirmaciones que entre sucesos. Sean dos sucesos A y B dados por afirmaciones válidas α y β. Entonces decir que A es independiente de B equivale decir que saber si α es cierta o no, no me permite tener nueva información sobre si β es cierta o no.

Para efectos de comprensión, es recomendable pensar que una afirmación válida α, más que dar un suceso A, da una partición de Ω, es decir, $\{A, A^c\}$. Esta partición mide la *cantidad de información* asociada a esta afirmación. Entonces, si dos afirmaciones α y β, tienen asociadas dos particiones $\{A, A^c\}$ y $\{B, B^c\}$, respectivamente; ambas afirmaciones en conjunto particionan el espacio en 4 sucesos $\{A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c\}$; es decir, 2 operaciones con 2 posibles resultados diferentes, me dan 4 resultados diferentes. La noción natural de independencia sería que, la "ganancia de información" debiera ser "máxima" al combinar ambas afirmaciones. Por ejemplo, si $A \subseteq B$ o $B \subseteq A$, la partición obtenida resulta ser una de las dos iniciales. O sea, no hay ganancia de información; Conocer una de las dos afirmaciones determina la otra. En este caso, los sucesos no son independientes por $\mathbb{P}$, sin importar cual sea la medida de probabilidad $\mathbb{P}$ que se tenga.

La definición de independencia es, entonces, global; es decir, entre las particiones $\{A, A^c\}$ y $\{B, B^c\}$. Por lo tanto, debería ser equivalente decir que A es independiente de B a decir que A es independiente de B^c o que A^c es independiente de B o que A^c es independiente de B^c.

En efecto, esto es cierto y basta mostrar que, si A es independiente de B entonces A^c es independiente de B para luego cambiar A por A^c y/o B por B^c. Este caso es bastante simple de probar:

$$\mathbb{P}(B) = \mathbb{P}(B \cap A^c) + \mathbb{P}(B \cap A) \Rightarrow \mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(B \cap A).$$

Además como A es independiente de B, por lo anterior:

$$\mathbb{P}(B \cap A^c) = \mathbb{P}(B)[1 - \mathbb{P}(A)] = \mathbb{P}(B)\mathbb{P}(A^c).$$

La relación entre sucesos de probabilidad positiva, dada por "ser independientes", no es refleja (a menos que $\mathbb{P}(A) = 1$) aunque si es simétrica. ¿Esta relación, es transitiva?. Es decir, dados A, B y C tres sucesos tales que A es independiente de B y B es independiente de C se tiene que ¿A es independiente de C?. No parece evidente si esto es cierto o no.

¿Qué ocurre si consideramos 3 sucesos A, B, C en vez de 2?, ¿cómo definir su independencia?. Es evidente que debieran ser independientes 2 a 2 (es decir tomando cualquier par de ellos). Por otro lado, también se debería cumplir que $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$. Por desgracia, es necesario pedir ambas propiedades, porque ninguna de ellas implica la otra. Esto se puede ver en el siguiente ejemplo:

Ejemplo 10. *Si consideramos* $\Omega = \{1, 2, 3, 4\}$ *con* $\mathbb{P}$ *equiprobable sobre* Ω*. Definimos* $A = \{1, 2\}$*,* $B = \{1, 3\}$ *y* $C = \{1, 4\}$*. Veremos que* A*,* B *y* C *son independientes 2 a 2 pero no se tiene que* $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$*.*

Respuesta: En efecto,

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A)\mathbb{P}(B), \\
\mathbb{P}(A \cap C) &= \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A)\mathbb{P}(C), \\
\mathbb{P}(B \cap C) &= \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(B)\mathbb{P}(C).
\end{aligned}$$

Sin embargo, $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{1\}) = \frac{1}{4}$ y $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{8}$.

Esto complica la definición de independencia, de una familia de sucesos cualquiera.

Definición 2. Familia de Sucesos Independientes. *Una familia* $\mathcal{C}$*, de sucesos, se dice* independiente por $\mathbb{P}$*, si para cualquier subconjunto finito* A_1*,...,*A_N *de* $\mathcal{C}$ *se tiene que:*

$$\mathbb{P}(A_1 \cap \cdots \cap A_N) = \mathbb{P}(A_1) \ldots \mathbb{P}(A_N)$$

La razón para considerar tan sólo intersecciones finitas en esta definición, quedará más clara más adelante. Tal como en el caso de dos sucesos, si una familia finita de sucesos es independiente, entonces cualquier combinación de complementos también lo será, es decir:

Lema 5.6. *Si* $\{A_1, \ldots, A_N\}$ *es una familia independiente por* $\mathbb{P}$ *entonces* $\{\bar{A_1}, \ldots, \bar{A_N}\}$ *también lo es, con:*

$$\bar{A_n} = \begin{cases} A_n \\ A_n^c \end{cases} .$$

Demostración. La demostración sigue las mismas líneas que el caso de dos sucesos. □

5.7 Probabilidades y el Infinito

En general hemos visto fórmulas para el cálculo de un número finito de elementos, con un número finito de operaciones (intersecciones, uniones, probabilidades condicionales, etc...). En espacios muestrales más que numerables vimos que es necesario realizar más que eso. Esto motiva extender algunas de estas fórmulas, a familias numerables de sucesos. Tal vez el ejemplo más sencillo y que resulta ser esencial, consiste en repetir una cantidad numerable de veces un experimento finito.

Veamos un ejemplo sencillo.

Ejemplo 11. *Se lanza una moneda con probabilidad de cara p, de manera independiente hasta que sale cara. ¿Cuál es la probabilidad de que se haya lanzado N veces?, ¿cuál es la probabilidad de que nunca salga cara?.*

El Experimento es $\mathcal{E}=$"Lanzar una moneda hasta que salga cara" y nos interesan las afirmaciones $\alpha_n=$"Salió cara en el n-ésimo lanzamiento" con $n \in \mathbb{N}$.

Como el experimento es lanzar una moneda, y en cada uno de los lanzamientos hay dos posibilidades distintas (cara y sello), un Espacio Muestral natural será de la forma $\Omega = \{0,1\}^{\mathbb{N}}$. Es decir el espacio formado por el producto del conjunto $\{0,1\}$ consigo mismo $\mathbb{N}$ veces.

Para cualquier $n \in \mathbb{N}$, α_n es una afirmación válida en este espacio muestral y por lo tanto tiene un suceso asociado A_n. El suceso definido por $\beta_N=$"la primera cara sale en el lanzamiento N-ésimo" es $B_1 = A_1$ y si $N > 1$ entonces $B_N = A_N \cap A^c_{N-1} \cdots \cap A^c_1$.

Siendo que la moneda tiene probabilidad de cara p y dado que se sigue lanzando la moneda, sólo si han salido sellos anteriormente tendremos por la fórmula 5.2 que, para cualquier $N > 1$:

$$\begin{aligned}
\mathbb{P}(B_N) &= \mathbb{P}(A_N \cap A^c_{N-1} \cdots \cap A^c_1) = \overbrace{\mathbb{P}(A_N | A^c_{N-1} \cap \cdots \cap A^c_1)}^{p} \cdot \\
&\cdot \overbrace{\mathbb{P}(A^c_{N-1} | A^c_{N-2} \cap \cdots \cap A^c_1)}^{1-p} \cdot \cdots \cdot \overbrace{\mathbb{P}(A^c_2 | A^c_1)}^{1-p} \cdot \overbrace{\mathbb{P}(A^c_1)}^{1-p} \\
&= p(1-p)^{N-1}.
\end{aligned}$$

Entonces, lanzar infinitas veces la moneda, es el complemento del suceso $C = \biguplus_{N \geq 1} B_N$. Pero,

$$\mathbb{P}(C) = \sum_{N \geq 1} \mathbb{P}(B_N) = \sum_{N \geq 1} p(1-p)^{N-1} = p \cdot \frac{1}{1-(1-p)} = 1.$$

Es decir, la probabilidad de lanzar infinitas veces la moneda es 0, lo que **NO** quiere decir que no ocurre.

Nota: ¿Es posible sacar el suceso C del espacio muestral?.

Veamos otro ejemplo.

Ejemplo 12. *Se tiene una moneda en que la probabilidad de obtener cara es* $p > \frac{1}{2}$*; un árbitro, en un partido, debe decidir cuál de los dos equipos tendrá el balón al comenzar el partido; La manera tradicional de decidirlo es lanzar una moneda y según* *el resultado un equipo o el otro tendrá el balón; El árbitro quiere ser justo y por lo tanto no quiere usar la moneda; El guardalinea sugiere que el árbitro use la moneda, lanzándola dos veces, y decidiendo como sigue; si sale cara y luego sello el primer equipo comienza, si sale sello y luego cara el segundo equipo comienza y si salen dos caras o dos sellos entonces la vuelve a lanzar dos veces y se repite el mismo razonamiento.*

¿ Es justo para ambos equipos?, ¿cuál es la probabilidad de que el arbitro nunca deje de lanzar la moneda?.

Un Espacio Muestral razonable sería $\Omega = \{-1, 0, 1\}^{\mathbb{N}}$, en que cada coordenada es una realización del experimento y -1 representa obtener sello-cara, 1 representa obtener cara-sello y 0 representa cara-cara o sello-sello. Nos interesan los sucesos A, B y C definidos por las afirmaciones válidas $\alpha =$"El primer equipo comienza"; $\beta =$"No se termina nunca de lanzar" y, $\gamma =$"El segundo equipo comienza". Es claro que A, B y C son disjuntos y que $\mathbb{P}(A \cup B \cup C) = 1$, entonces, el sistema será justo, si $\mathbb{P}(A) = \mathbb{P}(C)$, y siempre terminará si $\mathbb{P}(B) = 0$.

Para poder calcular las probabilidad de A, B y C, definimos algunos sucesos intermedios. Dado $n \in \mathbb{N}$, definimos las afirmaciones α_n="salió sello-cara la n-ésima vez", β_n="salió sello-sello o cara-cara la n-ésima vez" y, γ_n="salió cara-sello la n-ésima vez". Estas afirmaciones son válidas en nuestro Espacio Muestral y definen conjuntos A_n, B_n y C_n, respectivamente.

La probabilidad de obtener cara y luego sello en dos lanzamientos de una moneda con probabilidad de cara p, es igual a la probabilidad de obtener sello y luego cara en dos lanzamientos, es decir, $p(1-p)$. La probabilidad de obtener dos caras o dos sellos en dos lanzamientos será entonces $p^2 + (1-p)^2$. Por la propiedad 1 (y por álgebra elemental también), se tiene que $2p(1-p) = 1 - p^2 - (1-p)^2$. Para simplificar notemos como $q = p^2 + (1-p)^2$ y por lo tanto $p(1-p) = \dfrac{1-q}{2}$.

Lo que hay que condicionar, para poder hacer los cálculos, es el número de veces que hubo que lanzar la moneda dos veces. El suceso $E_N = \bigcap\limits_{n=1}^{N} B_n$ es el que asegura que se tuvieron que realizar al menos N veces los dos lanzamientos de la moneda. Por la fórmula 5.2 se tendrá que:

$$\mathbb{P}(E_N) = \prod_{n=2}^{N} \overbrace{\mathbb{P}(B_n | B_{n-1} \cap \cdots \cap B_1)}^{q} \cdot \overbrace{\mathbb{P}(B_1)}^{q} = q^N.$$

Es claro que $\mathbb{P}(A_N | E_{N-1}) = \mathbb{P}(C_N | E_{N-1}) = p(1-p)$ por lo que $\mathbb{P}(A_N \cap E_{N-1}) = \mathbb{P}(C_N \cap E_{N-1}) = q^{N-1} \cdot p(1-p)$.

Tenemos que $A = \biguplus\limits_{n=1}^{\infty} A_n \cap E_{n-1}$ y que $C = \biguplus\limits_{n=1}^{\infty} C_n \cap E_{n-1}$ es decir:

$$\mathbb{P}(C) = \mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(A_n \cap E_{n-1}) = \sum_{n=1}^{\infty} q^{N-1} \cdot p(1-p) = \frac{p(1-p)}{1-q} = \frac{1}{2}.$$

Entonces el sistema es justo y además $\mathbb{P}(B) = 1 - \mathbb{P}(A) - \mathbb{P}(C) = 0$, por lo que, con probabilidad 1 el árbitro dejará de lanzar la moneda.

- El suceso C no es vacío pero $\mathbb{P}(C) = 0$. ¿Entonces se podría sacar del espacio muestral el suceso C?.
- Desde un punto de vista teórico nunca verificamos que lo que definimos fuese una medida de probabilidad ni siquiera, que esto definiera una única medida de probabilidad.

Una de las dificultades del problema anterior es, que la afirmación "salió cara-sello la N-ésima vez" en castellano asume implícitamente la afirmación "no salió ni cara-sello ni sello-cara en las primeras $N-1$ veces" y a nivel de sucesos esto no es así. A pesar de que el Espacio Muestral era infinito, bastó considerar intersecciones y uniones numerables de sucesos, que dependían de un número finito de realizaciones del experimento para determinar la probabilidad de un suceso que dependía de todas las realizaciones (el suceso C).

En general, al considerar una cantidad numerable (a lo más) de repeticiones ,de un experimento, con Espacio Muestral finito este es siempre el caso.

En los dos ejemplos anteriores, el Espacio Muestral incluía la posibilidad de realizar infinitas veces un experimento con otro Espacio Muestral finito, pero la probabilidad de que eso ocurriera resultó nula en ambos casos. Como ya dijimos, no es claro que lo que definimos, para hacer los cálculos, fuese una medida de probabilidad, ni tampoco si era única. En este tipo de Experimentos veremos que los "singleton", no sirven de nada, para definir la medida de probabilidad.

Ejemplo 13. *Se lanza infinitas veces una moneda equilibrada. ¿Cuál es la probabilidad de obtener un resultado cualquiera fijo?.*

El Espacio Muestral natural en este caso sería $\Omega = \{0,1\}^{\mathbb{N}}$. Es claro que, si la moneda está equilibrada, entonces la probabilidad de cualquier resultado es la misma así que podemos calcular la probabilidad de obtener sólo caras.

Dado $n \in \mathbb{N}$, si definimos, como A_n, al suceso determinado por la afirmación "salió cara en el n-ésimo lanzamiento", entonces sabemos que $\mathbb{P}(A_n) = \dfrac{1}{2}$. Por otro lado, el suceso obtener cara en los primeros N lanzamientos, es $\bigcap\limits_{n=1}^{N} A_n$ y por la fórmula 5.2 se tendrá que:

$$\mathbb{P}(\bigcap_{n=1}^{N} A_n) = \overbrace{\mathbb{P}(A_N | \bigcap_{n=1}^{N-1} A_n)}^{\frac{1}{2}} \cdot \dots \cdot \overbrace{\mathbb{P}(A_2|A_1)}^{\frac{1}{2}} \cdot \overbrace{\mathbb{P}(A_1)}^{\frac{1}{2}} = \frac{1}{2}^N.$$

El suceso obtener sólo caras, se escribe como $A = \bigcap_{n \in \mathbb{N}} A_n$ y dado que $A \subseteq \bigcap_{n=1}^{N} A_n$ para cualquier $N \in \mathbb{N}$ por la propiedad 5.3.2 se tendrá que $\mathbb{P}(A) = 0$.

Entonces, la probabilidad de obtener cualquier secuencia de resultados es 0; por lo tanto, la probabilidad de obtener cualquier resultado, de todos los posibles, es 0; esto, en particular, muestra que definir la probabilidad sobre los singleton, no caracteriza en lo más mínimo la función de probabilidad, cuando el espacio es más que numerable.

5.8 Ejercicios

1. Dos jugadores A y B juegan lanzando cada uno dos dados simultáneamente; A gana si obtiene una suma 6 de los dos dados, antes que B saque suma 7 y, B gana si saca suma 7 antes que A saque suma 6. Muestre que la probabilidad de que gane A y la probabilidad de que gane B están en proporción 30 : 31. ¿Cuál es la probabilidad de que no gane ninguno de los dos?.

2. Tres jugadores A, B y C de igual habilidad se enfrentan en un campeonato que consiste en una serie de juegos en que A juega con B, B juega con C, C juega con A, A juega con B, y así sucesivamente en rotación. El jugador que gana dos juegos consecutivos gana el campeonato. Muestre que las probabilidades de ganar de A, B y C están en proporción 12 : 20 : 17. ¿Cuál es la probabilidad de que no gane ninguno de los tres?.

3. A y B juegan ,uno contra el otro, con dos dados; A gana si obtiene 6 puntos, y B gana si obtiene 7 puntos; le corresponde el primer tiro a A, los dos siguientes a B, los otros dos siguientes a A, y así sucesivamente, hasta que gane gane alguno de los dos jugadores. La pregunta es: ¿Cuál es la probabilidad de A sobre B?
   Nota: Propuesto por Fermat en carta a Huygens en junio de 1656, y resuelto por este en carta a Carcavi el mes siguiente.

4. Tres jugadores A, B y C, teniendo 12 fichas de las cuales cuatro son blancas y ocho negras, juegan con la condición de que gana el primer jugador que obtiene (al extraer sin mirar) una ficha blanca; A extrae primero, luego B, luego C, luego A nuevamente, y así sucesivamente. ¿Cuál es la proporción de las probabilidades de ganar de cada jugador con respecto de los otros?
   Respuesta: 9:6:4
   Nota: Resuelto por Huygens en 1665
5. Como antes, los jugadores tienen 12 fichas de las cuales cuatro son blancas y ocho negras; A apuesta a B, que escogiendo siete fichas sin mirar, obtendrá tres blancas. ¿Cuáles son las probabilidades de A sobre B?

   (Huygens también considera el caso en que A apuesta a escoger tres o mas fichas blancas.)
   Nota: Resuelto por Huygens en 1665.

6. Teniendo A y B cada uno doce fichas, juegan con tres dados, bajo la condición de que si se obtienen 11 puntos, A entrega una ficha a B, si se obtienen 14, B entrega una a A, ganando el jugador que obtiene primero todas las fichas.¿Cuáles son las probabilidad de A sobre B?
   Nota: Problema planteado por Pascal a Fermat y, a través de Carcavi, a Huygens en septiembre de 1656, en una carta que contiene las soluciones dadas por Pascal y Fermat. La solución de Huygens está en una carta, a Carcavi, un mes después y, la demostración en una nota de 1676. Este problema se conoce como el *Problema de la Ruina del Jugador*.

Bibliografía

[1] Cardano, G, *The book on games of chance*, traducido por Gould S. H., Rinehart and Winston, Nueva York

[2] Degroot M. H., *Probabilidad y Estadísitica*, Addison Wesley

[3] Feller, W, *Introducción a la Teoría de Probabilidades y sus aplicaciones*, ISBN 968-18-0721-9, Grupo Noriega Editores

[4] Kolgomorov, A. N., *Fundaciones de la Teoría de Probabilidades*, 1953, Chelsea Publishing Company, Nueva York

[5] Meyer, P., *Probabilidad y aplicaciones estadísticas*, Addison Wesley

[6] Neveu, J., *Bases Mathématiques du Calcul des Probabilités*, Masson et Cie, 1970, París, Francia

[7] Basulto, *La geometría del azar: La correspondencia entre Pierre de Fermat y Blaise Pascal*, Nibola Libros Ediciones, ISBN 9788496566545, Madrid, España

[8] Santos del Cerro, J et Al, *Historia de la Probabilidad y Estadística (III)*, Delta, ISBN: 9788496477254, Madrid, España

Índice de figuras

Índice de términos