UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

TEORÍA DE LA PROBABILIDAD PARA LA ESTADÍSTICA MATEMÁTICA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICA

P R E S E N T A :

ALEJANDRA CALVO FLORES

DIRECTOR DE TESIS:
DR. SERGIO HERNÁNDEZ CASTAÑEDA
2010

1. Datos del alumno Calvo Flores Alejandra 11 14 87 08 Universidad Nacional Autonoma de Mexico Facultad de Ciencias Matematica 09653469
2. Datos del tutor Dr Sergio Hernandez Castañeda
3. Datos del sinodal 1 M en C Maria de la Paloma Carmen Zapata Lillo
4. Datos del sinodal 2 M en A P Maria del Pilar Alonso Reyes
5. Datos del sinodal 3 Dra Guadalupe Carrasco Licea
6. Datos del sinodal 4 M en C Jose Antonio Flores Diaz
3. Datos de la tesis Teoria de la Probabilidad para la Estadistica Matematica 115 p 2010

Esta tesis representa el fin de una etapa muy enriquecedora en mi vida y está dedicada:

A mis padres Esperanza y Alejandro.
Gracias por su apoyo incondicional, su guía y su confianza en la realización de mis sueños. Soy afortunada por contar siempre con su amor, comprensión y ejemplo. Esta tesis es suya.

A Paco y a Sofía.
Gracias por hacer mi vida tan maravillosa, por su comprensión, amor, paciencia y por estar siempre a mi lado.

***A Luis** con mucho cariño, sabes que cuentas conmigo siempre.*

Agradecimientos

Toda esta temática no fue fácil de abordar, pero durante el desarrollo de este trabajo tuve la fortuna de recibir valiosos comentarios, sugerencias y un gran apoyo del grupo de economía matemática y teoría de juegos de la Facultad de Ciencias, cuyos nombres enlisto en orden alfabético Salvador Ferrer, Rubí Gutiérrez por sus asesorías con Matlab, Alma Jiménez, César Sousa, Francisco Valdez, Claudia Villegas y por supuesto sus coordinadores, Paloma Zapata y Sergio Hernández; todos ellos participaron directa o indirectamente en todo momento.

También quiero agradecer a mis sinodales quienes tuvieron la paciencia de revisar y hacer observaciones para mejorar mi trabajo, en especial agradezco a la profesora Guadalupe Carrasco.

Al profesor Sergio Hernández Castañeda mi más sincero agradecimiento, por su paciencia y apoyo brindados desde que empecé con esto, por la vocación y entusiasmo con las que hace su trabajo, por ser un gran ejemplo y además por ser el principal inspirador del gusto que tengo por la economía.

Gracias a todos.

Alejandra Calvo Flores

Índice general

Introducción

En este trabajo se desarrolla una introducción a la teoría de la probabilidad elemental y a la estadística matemática, con el objetivo de ofrecer material que resulte útil a estudiantes y en especial a profesores de las carreras de economía. Ésta temática puede ser interesante para los economistas, debido a que algunas de éstas técnicas son empleadas en el análisis de fenómenos económicos bajo incertidumbre y en econometría.

Al principio, una de las ideas que originaron este trabajo se basa en lo siguiente: se sabe que para poder entender a fondo la teoría de la probabilidad se requiere de un acercamiento a la teoría axiomática de Kolmogórov que a su vez exige una aproximación a la teoría de la medida, rama de las matemáticas bastante sofisticada. Asumiendo lo anterior, e incluyendo el propósito de ser pedagógicos y, por supuesto, no exigir a los alumnos ningún tipo de contacto con la teoría de la medida, se propone que una buena ruta para llevar a cabo lo señalado puede consistir en centrar un curso en el estudio de los espacios discretos de probabilidad y, con base en ellos, introducir conceptos elementales y, en general, los conceptos fundamentales de la teoría.

Así fue concebido este trabajo y se distribuye en cinco capítulos los cuales se detallan a continuación.

El contenido del primer capítulo está inspirado en el libro de Gnedenko, es preciso señalar que aquí no empieza el desarrollo de la propuesta de este trabajo, en este primer capítulo se hace una introducción en donde se resalta la importancia del enfoque axiomático y en consecuencia el por qué del uso de los espacios finitos de probabilidad. En él se abordan las tres diferentes aproximaciones a la definición de probabilidad, clásica, frecuentista y axiomática; se sigue el proceso de cómo surgieron a lo largo de la historia, y se presentan algunos de los posibles problemas a los que se pueden enfrentar los alumnos a la hora de aprender los conceptos de la teoría de la probabilidad.

En el segundo capítulo, se desarrollan los conceptos básicos de la teoría de la probabilidad como los axiomas, los espacios finitos de probabilidad y algunos conceptos elementales, como probabilidad condicional, independencia de eventos y los espacios producto de probabilidad.

El tercer capítulo, está dedicado básicamente al significado del teorema de DeMoivre-Laplace que viene motivado por los experimentos Bernoulli. Es preciso comentar que se tuvo especial cuidado en la demostración de este teorema, en la búsqueda de que conceptos como puede ser en este caso, el de convergencia uniforme sea captado del mejor modo posible por el estudiante. Se incluye también la versión integral del teorema DeMoivre-Laplace con su demostración y un resultado relativo a la ley de los grandes números.

En el cuarto capítulo, se aborda el concepto de variable aleatoria y sus características numéricas. Por último, en el quinto capítulo se discute una de las versiones del teorema central del límite, además se exhibe su relación con el teorema integral de DeMoivre-Laplace. En las aplicaciones de este capítulo se incluye un programa elaborado en Matlab en donde se generaliza el experimento de la moneda, ahora se hace con una figura de tres lados, después con un dado y con lo anterior se puede ver cómo podría ser la generalización para el número de caras que se desee.

Finalmente en el apéndice A se presentan algunos conceptos básicos relacionados con la notación y las operaciones booleanas, se define de manera general el concepto de función y algunas definiciones relacionadas.

ACF 2010

Capítulo 1

Breve contexto histórico

En mi experiencia dando clases he percibido, que los estudiantes encuentran los resultados de probabilidad y estadística poco confiables, vagos y difíciles. Estas dificultades persisten, por la inadecuada definición de los conceptos fundamentales, dando lugar a una confusión constante entre lo que se intuye y las conclusiones lógicas.

En este capítulo, se presentan los tres diferentes enfoques para aproximarse a la definición de probabilidad, enfoque clásico, enfoque frecuentista y enfoque axiomático. Se sigue el proceso de cómo surgieron a lo largo de la historia con el objetivo de presentar algunos de los posibles problemas a los que se pueden enfrentar los alumnos a la hora de aprender los conceptos básicos de la teoría de la probabilidad. Así pues, conociendo las ventajas y limitaciones que ofrece cada una de las tres diferentes aproximaciones, se tiene un panorama para justificar lo que se propone en este trabajo. Pues la idea consiste en centrar un curso basado en los espacios finitos de probabilidad que se desprenden como consecuencia de la construcción axiomática.

Conceptos básicos

La teoría de la probabilidad estudia experimentos cuyo resultado no puede determinarse antes de realizarlos. Por ejemplo, antes de lanzar un dado no es posible determinar qué número se obtendrá, tampoco se puede saber si al lanzar una moneda saldrá águila o sol, o al barajear las cartas no sabemos qué cartas nos saldrán. A este tipo de experimentos se les conoce como aleatorios.

Cuando se decide modelar un experimento ya sea real o imaginario lo primero que se hace es decidir cuales son los posibles resultados de tal experimento, a este conjunto le llamaremos *espacio muestra* y se denota con Ω.

El *espacio muestra* Ω, también es conocido como el *evento cierto*, cada uno de los elementos que lo forma es llamado *evento elemental*, así mismo el conjunto vacio $\varnothing$ es el *evento imposible* y los *eventos* son subconjuntos de Ω.

Ejemplo 1 *Se lanza una moneda. El espacio muestra es* $\Omega = \{\acute{a}guila, sol\}$ *el cual contiene dos eventos elementales águila y sol, y aunque el conjunto vacio* $\varnothing$ *también pertenece a* Ω, *este no es considerado como un evento elemental.*

Ejemplo 2 *Se lanza un dado. Sea* $\Omega = \{1, 2, 3, 4, 5, 6\}$ *el conjunto* Ω *tiene 6 eventos elementales.*

Si A es el evento la cara es par, entonces

$$A = \{2, 4, 6\}.$$

Si B es el evento la cara es menor que 3, entonces

$$B = \{1, 2\}$$

A y B son subconjuntos de Ω.

Enfoque clásico

El enfoque se asocia principalmente a Fermat, Pascal, J. Bernoulli a mediados del siglo XVII.

A continuación se presenta una primera aproximación a la definición clásica en la cual el significado de probabilidad se reduce a un concepto todavía más sencillo, el de *equiprobable*, es decir que todos los eventos tienen la misma probabilidad de ocurrir.

Definición *Dado un experimento la probabilidad del evento* A *es igual a los casos favorables entre los casos totales, es decir*

$$P(A) = \frac{\text{Número de casos favorables}}{\text{Número de casos totales}}$$

Ejemplos

1. *En el exprimento de lanzar un dado honesto.* Es fácil determinar que los casos totales son 6 y si el evento es obtener un número impar entonces los casos favorables son 3. Así la probabilidad de obtener un número impar es:

   $$P(impar) = \frac{3}{6}.$$

   Como el dado es honesto sus caras tienen la misma probabilidad de salir, entonces sucede que:

   $$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6},$$

en consecuencia la probabilidad de obtener un número impar es

$$\begin{aligned} P(impar) &= P(1) + P(3) + P(5) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6}. \end{aligned}$$

Es importante notar que no siempre es tan claro identificar quienes son los casos favorables y contar quienes son los casos totales. Con el siguiente ejemplo se resaltaran algunas de las posibles ambigüedades en esta definición.

2. *Se lanzan dos dados y el evento X es la suma de sus caras da siete.*

   Supongamos que el primer dado lanzado es rojo y el segundo azul, se podría pensar que se tienen 21 casos totales que serían los siguientes:

$$\left\{ \begin{array}{cccccc} (1-1) & (1-2) & (1-3) & (1-4) & (1-5) & (\mathbf{1}-\mathbf{6}) \\ & (2-2) & (2-3) & (2-4) & (\mathbf{2}-\mathbf{5}) & (2-6) \\ & & (3-3) & (\mathbf{3}-\mathbf{4}) & (3-5) & (3-6) \\ & & & (4-4) & (4-5) & (4-6) \\ & & & & (5-5) & (5-6) \\ & & & & & (6-6) \end{array} \right\}$$

De los cuales se puede ver que sólo tres casos suman 7 por lo tanto usando la definición clásica se tiene que:

$$P(X) = \frac{3}{21}.$$

Ahora supóngase que se lanzan los mismos dados rojo y azul, pero ahora tenemos 36 casos totales

$$\left\{ \begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (\mathbf{1},\mathbf{6}) \\ (2,1) & (2,2) & (2,3) & (2,4) & (\mathbf{2},\mathbf{5}) & (2,6) \\ (3,1) & (3,2) & (3,3) & (\mathbf{3},\mathbf{4}) & (3,5) & (3,6) \\ (4,1) & (4,2) & (\mathbf{4},\mathbf{3}) & (4,4) & (4,5) & (4,6) \\ (5,1) & (\mathbf{5},\mathbf{2}) & (5,3) & (5,4) & (5,5) & (5,6) \\ (\mathbf{6},\mathbf{1}) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array} \right\}$$

y los casos favorables son 6 por lo tanto siguiendo la definición clásica se tiene que:

$$P(X) = \frac{6}{36} = \frac{1}{6}.$$

Como se puede notar la definición clásica no es tan clara, ya que en ambos casos se usa correctamente y sin embargo se obtienen resultados distintos. Por lo que la definición clásica podría ser cuestionada.

Lo que sucede con el ejemplo anterior es lo siguiente:

Se dice que un espacio muestral es *equiprobable* si no hay algún punto muestral que sea más favorecido por el experimento que otros puntos muestrales; es decir, que no haya algo en el experimento que favorezca la ocurrencia de un punto muestral respecto a otros. Nótese que en el primer espacio muestral, donde se tienen 21 casos totales, el punto muestral formado por los números 1 y

3 puede ocurrir de dos maneras distintas (1 en el dado rojo y 3 en el azul o bien 1 en el dado azul y 3 en el rojo) mientras que el formado por dos veces el 2 sólo puede ocurrir de una manera (2 en cada dado). Por tanto no es un espacio muestral equiprobable. La notación presentada es una forma de escribir el espacio muestral y puede usarse para calcular probabilidades, pero se requiere tener en cuenta que sus elementos pesan diferente y se debe contar adecuadamente. En cambio en el caso donde tenmos 36 casos totales el espacio es equiprobable y en ese caso se pueden calcular probabilidades aplicando directamente la definición clásica.

Éste no es el único problema que pudiera presentarse, por lo que a continuación se enlistan algunas de las posibles debilidades de la definición clásica, cabe destacar que no por ello dejará de tener importancia.

Problemas con la definición clásica.

- Es difícil determinar quienes son los casos favorables y casos totales, como se mostró en el ejemplo anterior.

- Sólo se puede aplicar a los casos en donde los eventos son equiprobables. Generalmente los casos donde son equiprobables son los juegos de azar.

- Esta definición está totalmente separada de la experiencia.

  Por ejemplo, si se desea determinar la probabilidad de que al nacer un bebé sea niño, se asume generalmente que $P(niño) = \frac{1}{2}$, pero esto es sólo un resultado aproximado, pues no se estan tomando en cuenta las características familiares del bebé, ni el lugar de nacimiento, ni su carga genética.

- Por último, no considera cuando los casos favorables y casos totales son infinitos.

  Por ejemplo, si se desea determinar ¿cuál es la probabilidad de que se gane un sueldo de 10,000 pesos cuando se termina una licenicatura? En realidad la cantidad de salarios posibles, es un número infinito que podría considerarse en la recta real positiva pues, es factible que se pueda ganar un número en el intervalo $[0, \infty)$. Entonces $P(10,000) = 0$, ya que los casos totales son infinitos.

Probabilidad Geométrica

En el último punto de la página anterior se hizo referencia al problema cuando se tienen casos favorables y casos totales infinitos. Al respecto se puede decir que cuando se popularizó la definición clásica de probabilidad, surgieron ideas acerca de cómo generalizar esta definición a experimentos aleatorios en los cuales el espacio muestral es infinito no numerable. En estos casos no es posible recurrir al conteo de los casos favorables ni de los casos totales, como en los ejemplos usuales de los dados o monedas, pero si es posible usar medidas geométricas.

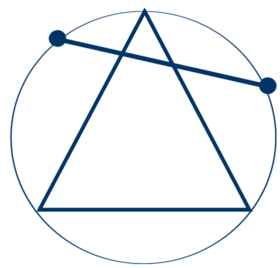
Para calcular la probabilidad de un evento A de este estilo se puede usar una medida como el área, entonces por definición

$$P(A) = \frac{\text{área de la región que representa } A}{\text{área total}} \tag{1.1}$$

que es la probabilidad de que un punto escogido aleatoriamente en la región del aréa total caiga en la región A. Ahora se plantea un problema clásico, que se conoce como la paradoja de Bertrand, en el cual podremos notar las dificultades que existirían al utilizar áreas, longitudes o alguna otra medida en la definición clásica.

La paradoja de Bertrand

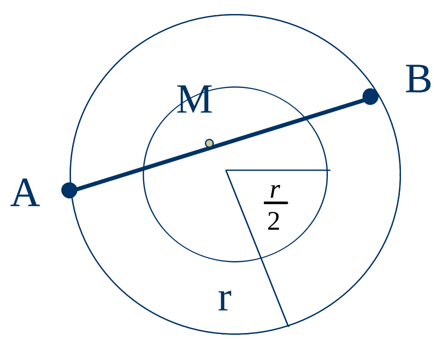
Se elige aleatoriamente una cuerda en un círculo. ¿Cuál es la probabilidad de que su tamaño exceda la medida del lado de un triángulo equilátero inscrito? Ver la siguiente figura:



CASO 1. Si M es el punto medio de la cuerda AB cae dentro del círculo de radio $\frac{r}{2}$, entonces la cuerda es más grande que el lado del triángulo. Por lo tanto los casos totales son todos los puntos dentro del circulo de radio r, y los casos favorables son los puntos dentro del círculo de radio $\frac{r}{2}$. Usando la definición 1.1 se tiene que:

$$P = \frac{\pi\left(\frac{r}{2}\right)^2}{\pi r^2} = \frac{1}{4}.$$

El numerador representa el área del círculo con radio $\frac{r}{2}$ y el denominador el área del círculo con radio r.

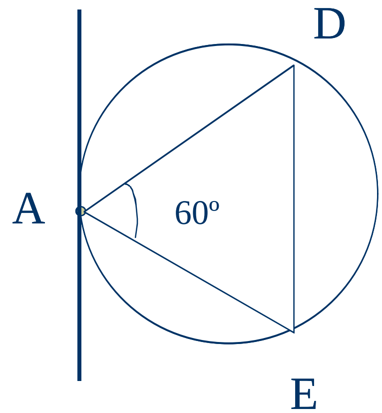


Paradoja de Bertrand (Caso1)

CASO 2. Se asume que A es un punto fijo sobre la circunferencia. La tangente que se encuentra en el punto fijo, y los lados del triángulo equilátero con vértice en este punto forman tres ángulos de 60^o. Los casos favorables son todas las rectas que se encuentran en el arco DE. Usando nuevamente la definición 1.1, se tiene lo siguiente:

$$P = \frac{\frac{2\pi r}{3}}{2\pi r} = \frac{1}{3}$$

El numerador representa $\frac{1}{3}$ del perímetro de la circunferencia y el denominador representa el perímetro de la circunferencia.



Paradoja de Bertrand (Caso 2)

Se han encontrado dos soluciones diferentes a este problema, con esto se hacen evidentes las ambigüedades en la definición clásica, en el caso en que los casos favorables y los casos totales son infinitos.

En realidad lo que sucede es que las soluciones pertenecen a dos problemas diferentes pues se trata de dos $\sigma - \acute{a}\lg ebras$ distintas.

Enfoque frecuentista

El enfoque se asocia principalmente a Laplace, R. Von Mises a mediados del siglo XIX.

La experiencia demuestra que muchos de los fenómenos no determinísticos exhiben una regularidad estadística que los hace objeto de estudio. La definición frecuentista toma como punto de partida la *frecuencia relativa* de la ocurrencia de un evento, en un fenómeno observado un número grande de intentos o ensayos.

Lo anterior se refiere a que otra forma de determinar la facilidad con la que ocurre un evento consiste en basarse en la experiencia práctica, repitiendo muchas veces el experimento y observando cuantas veces ocurre el evento. Pero el número de veces que ocurre el evento de interés varía al cambiar el número de veces que se repite el experimento, asi que lo importante es analizar la *frecuencia relativa* del evento, es decir la razón

$$P\left(A\right) = \frac{\text{Número de veces que ocurre el evento}}{\text{Número de pruebas realizadas}}.$$

Con el calculo anterior se puede ver que la probabilidad del evento A tiende a estabilizarse al rededor de un número.

Esto puede ser ilustrado con el siguiente ejemplo donde la frecuencia relativa es prácticamente una constante cuando el experimento se realiza un número grande de veces.

Ejemplo *Se desea obtener la probabilidad de que al lanzar un dado salga 2.*

Sea $A = \{2\}$, con la definición clásica sabemos que $P(A) = \frac{1}{6} \simeq 1{,}66\overline{6}$.

Se realizó el experimento de lanzar un dado, se calculó su probabilidad frecuentista variando el número de lanzamientos del dado y se obtuvo lo que a continuación se muestra en la tabla:

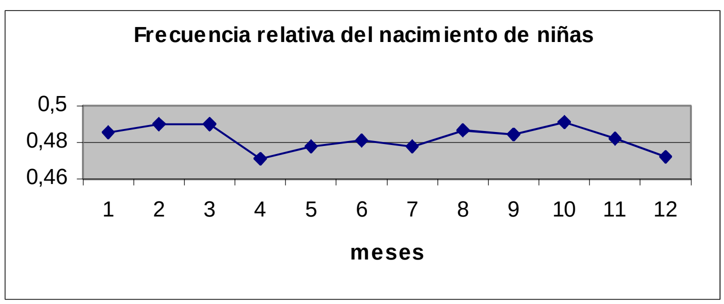
Número de lanzamientos	Probabilidad frecuentista
6	$\frac{2}{6} \simeq 0{,}333$
100	$\frac{18}{100} \simeq 0{,}180$
1000	$\frac{158}{1000} \simeq 0{,}158$

Con este sencillo ejemplo puede notarse que a medida que el número de lanzamientos aumenta, la probabilidad frecuentista se va acercando a la probabilidad clásica. Es preciso señalar, que los aficionados a los juegos de azar han usado este método desde hace siglos.

A continuación se presenta un ejemplo clásico.

Ejemplo *Se considera la distribución de recien nacidos de acuerdo al sexo y mes de nacimiento, en un año*[1].

Mes	1	2	3	4	5	6	7	8	9	10	11	12	Total
Nacimientos	7280	6957	7883	7884	7892	7609	7585	7393	7203	6903	6552	7132	88273
Niños	3743	3550	4017	4173	4117	3944	3964	3797	3712	3512	3392	3761	45682
Niñas	3537	3407	3866	3711	3775	3665	3621	3596	3491	3391	3160	3371	42591
Frec. Relativa niñas	0.486	0.49	0.49	0.471	0.478	0.482	0.477	0.486	0.485	0.491	0.482	0.473	0.482491815



Es decir, en el primer mes hubo 7280 nacimientos de los cuales 3743 fueron niños y 3537 niñas.

Por lo tanto, la frecuencia relativa de las niñas en el primer mes corresponde a la división de 3537 entre el total de nacidos en ese mes que son 7280 lo que resulta 0,486. La grafica muestra los resultados de todo un año, y se puede notar que la frecuencia relativa de las niñas fluctúa alrededor del número 0,482. En todos los casos donde se aplica la definición frecuencial, la frecuencia relativa toma lugar en una vecindad de la probabilidad P del evento, en este caso 0,482 es una vecindad de 0,5.

Es muy natural que la teoría de la probabilidad se aplique en muchos problemas que tienen que ver con el mundo real, pero se hace hincapíe que la definición frecuentista se basa sólo en hechos empíricos.

[1] Los datos para realizar la gráfica, fueron obtenidos de la Tabla 3 Gnedenko pag. 46.

Cabe resaltar que una vez que el experimento se realiza un número grande de veces lleva a la deducción de ciertas reglas, independientemente del caso que se este estudiando.

Un evento A tiene probabilidad $P(A)$ si este evento tiene las siguentes características:

- Es posible al menos en principio, realizarlo un número ilimitado de veces bajo las mismas condiciones, y en los cuales el evento puede o no puede ocurrir.

- Al realizarlo un número suficiente de veces, la probabilidad $P(A)$ del evento tiende a estabilizarse al rededor de su frecuencia relativa.

A partir de que se dió a conocer la definición frecuentista se fueron acumulando gran cantidad de datos con fines estadísticos y demográficos que reflejan resultados con mucha exactitud, a continuación se enlistan los problemas con esta definición.

Problemas con la definición frecuentista

- Es importante aclarar que las propiedades fueron obtenidas a partir de observaciones y no de una manera formal de carácter matemático.

- No se puede hablar de probabilidad de un evento hasta realizarlo un número ilimitado de veces, además de exigir su realización bajo las mismas condiciones.

Enfoque axiomático

El enfoque se asocia principalmente a Kolmogorov, aunque se consideran importantes otros autores como Bernstein, Borel, Khintchine aproximadamente por 1931.

El enfoque axiomático se desarrolla como una maduración de ideas que propician los otros enfoques y sirve para plantear la probabilidad de eventos que no se ajustan a las condiciones de éstos últimos, pero que a pesar de todo son de interés probabilístico, como podrían ser los fenómenos naturales.

Hasta ahora la teoría de la probabilidad no ha sido formulada matemáticamente. Esta falta de claridad permite conclusiones paradójicas como la de Bertrand. Y como era de esperarse todas las aplicaciones que se dieron de los fenómenos naturales carecían de solidez, que muchas veces originó críticas. El desarrollo de la ciencia a principios del siglo XX motivó el estudio de los conceptos fundamentales de la teoría de una manera sistemática, así mismo surgió la necesidad de esclarecer las condiciones bajo las cuales era posible usar los resultados de la teoría de la probabilidad, ésta fue una de las causas por la cual la construcción axiomática adquirió mucha importancia. Además las probabilidades usadas como se usan en los problemas reales son compatibles con los axiomas.

El acercamiento axiomático propuesto por Kolmogorov, relaciona a la teoría de la probabilidad con la teoría de conjuntos y la teoría de la medida ambos aspectos de la teoría de funciones de variable real. A continuación se presenta una breve descripción de la construcción axiomática.

Construcción axiomática

En esta sección se define un espacio en el sentido que lo hizo Kolmogorov. Para ello se enuncian algunas definiciones que de antemano reconozco que son complicadas, es importante aclarar que yo no pretendo que los alumnos aprendan teoría de la medida. Pero lo enuncio por dos razones, en primer lugar porque quisiera crear conciencia en los profesores de que detrás de la definición axiomática de probabilidad hay una gran sustento matemático y en segundo lugar porque mi trabajo propone una forma de evitar en la medida de lo posible este tipo de conceptos.

Para de definir un espacio de probabilidad, se ponen a consideración las siguientes definiciones.

Definición 1 *Sea Ω un conjunto. Se dice que una familia $\Im$ de subconjuntos de Ω es una* ***álgebra*** *si se satisfacen las siguientes condiciones:*

1. $\Omega \in \Im$.

2. *Si $A \in \Im$, entonces $A^c \in \Im$.*

3. *Si $A_1, A_2, ..., A_n$ es cualquier familia finita de elementos de $\Im$, entonces $\cup_{i=1}^{n} A_i \in \Im$.*

Es decir una álgebra es cerrada bajo uniones y también lo es bajo intersecciones y diferencias.

Definición 2 *Sea Ω un conjunto. Se dice que una familia $\Im$ de subconjuntos de Ω es una* $\boldsymbol{\sigma -}$ **álgebra** *si es una álgebra y dada cualquier familia infinita numerable de elementos de $\Im$, $A_1, A_2, ...,$ entonces $\bigcup_{i=1}^{\infty} A_i \in \Im$.*

Definición 3 *Sea Ω un conjunto y $\Im$ un álgebra de subconjuntos de Ω. Se dice que una función no negativa P, definida sobre $\Im$, es* ***finitamente aditiva*** *si dada cualquier familia finita, $A_1, A_2, ..., A_n$, de elementos de $\Im$, ajenos por parejas, entonces:*

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

Definición 4 *Sea Ω un conjunto y $\Im$ un álgebra de subconjuntos de Ω. Se dice que una función no negativa P, definida sobre $\Im$, es* $\boldsymbol{\sigma -}$ ***aditiva*** *(o que tiene la propiedad de la aditividad numerable) si es finitamente aditiva y dada cualquier familia infinita numerable de $A_1, A_2, ...,$ de elementos de $\Im$, ajenos por parejas, tales que $\bigcup_{i=1}^{\infty} A_i \in \Im$, entonces:*

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

Definición 5 *Sea Ω un conjunto y $\Im$ un álgebra de subconjuntos de Ω. Se dice que una función no negativa P, definida sobre $\Im$, es una* ***medida de probabilidad*** *si es $\sigma -$ aditiva y $P(\Omega) = 1$.*

Espacios de probabilidad en el sentido de Kolmogorov

Un espacio de probabilidad en el sentido de Kolmogorov es una terna $[\Omega, F, P]$, donde
Ω es un conjunto llamado conjunto de eventos elementales,
F una $\sigma-$álgebra de subconjuntos con unidad Ω,
P es una medida de probabilidad.

Y como se puede observar, no es tan pedagógico como se quisiera, pues se encuentran una serie de definiciones que por sí solas cuestan trabajo.

Características del enfoque axiomático

- Incluye como casos particulares a los enfoques clásico y frecuentista y supera los defectos que ambas tenían.
- Enfatiza el carácter deductivo de la probabilidad.
- Evita ambigüedades conceptuales.
- Ofrece una preparación sólida y sofisticada.
- Es la base para un estudio más profundo en esta rama.

Con las características señaladas se puede ver que el enfoque axiomático es el más preciso y esto da pie a pensar que es uno de los mejores caminos para introducir la teoría de la probabilidad, incluso si es un curso elemental y en realidad es un buen camino pero para los matemáticos.

Este camino no es el mejor si se trata de ser pedagógico, por esta razón la alternativa que se propone en este trabajo es que los alumnos aprendan a moverse en espacios finitos de probabilidad, los cuales estarán fundamentados en el enfoque axiomático y con ellos se simplificarán varias cosas.

En los capítulos subsecuentes empezaría lo que se considera la propuesta del curso.

Capítulo 2

Conceptos básicos

En el capítulo anterior, se presentaron algunos de los problemas a los que los estudiantes se enfrentan cuando estudian la teoría de la probabilidad, y también se señaló que uno de los caminos para eliminar estas dificultades es la construcción axiomática. Lo anterior da pie a pensar que la construcción axiomática es la mejor para introducir la teoría de la probabilidad, incluso si es un curso elemental, pero también puede notarse que es complicado.

En este capítulo se hace una aproximación a la construcción axiomática pero se desea evitar un lenguaje complicado, además de simplificar la notación y no propiciar las dificultades que se generan con los espacios en el sentido de Kolmogorov.

La alternativa que se propone es que los alumnos aprendan a moverse en espacios finitos de probabilidad. Como se podrá ver más adelante, los espacios finitos de probabilidad son simples ya que las probabilidades de los eventos quedarán expresadas en términos de los eventos elementales. Con ellos será posible abarcar conceptos básicos y en general, los conceptos fundamentales de la teoría y por supuesto no exigir a los alumnos ningún tipo de acercamiento a la teoría de la medida.

En este capítulo también se exponen los axiomas de probabilidad, la probabilidad condicional, el teorema de multiplicación, teorema de probabilidad total, el teorema de Bayes, independencia entre eventos y los espacios producto de probabilidad.

Espacios finitos de probabilidad

Sea $[\Omega, p]$ un sistema en donde Ω es un conjunto finito y $p : \Omega \to R$ tal que:

1. $p(a_i) \geqq 0$ para todo $a_i \in \Omega$
2. $\displaystyle\sum_{i=1}^{n} p(a_i) = 1.$

Un sistema como el considerado induce de forma unívoca un espacio de probabilidad en el sentido de Kolmogorov $[\Omega, F, P]$ (ver página 15), en donde naturalmente F en los espacios finitos sería 2^{Ω}, y si $A \in F$ el valor de $P(A)$, se define como la siguiente suma:

$$P(A) = \sum_{a_i \in A} p(a_i),$$

en donde la probabilidad del evento A queda expresada en términos de las probabilidades de los eventos elementales [1].

Ejemplos

Dentro de los espacios de probabilidad finitos se encuentran los espacios de probabilidad finitos igualmente probables y los que no son igualmente probables, para ilustrar ambos tipos se muestran los siguientes ejemplos.

Ejemplo 3 ***Lanzar una moneda honesta.*** *Sea $\Omega = \{\acute{a}guila, sol\}$ y para asociar las probabilidades se debe cumplir que*

$$\begin{aligned} p(\acute{a}guila) &\geq 0 \\ p\,(sol) &\geq 0, \end{aligned}$$

y además

$$P(\Omega) = p(\acute{a}guila) + p\,(sol) = 1,$$

la moneda es honesta, entonces se sabe que

$$\begin{aligned} p(\acute{a}guila) &= \frac{1}{2} \\ p(sol) &= \frac{1}{2}. \end{aligned}$$

Ejemplo 4 ***Lanzar una moneda cargada.*** *Sea $\Omega = \{\acute{a}guila, sol\}$. En una moneda cargada se da mayor probabilidad a uno de los dos eventos elementales, en este caso se asigna mayor probabilidad al evento águila sin descuidar que se debe cumplir que $p(\acute{a}guila) + p\,(sol) = 1$, por tanto:*

$$\begin{aligned} p(\acute{a}guila) &= \frac{2}{3} \\ p(sol) &= \frac{1}{3} \end{aligned}$$

Ejemplo 5 ***Se lanza un dado cargado.*** *Sea $\Omega = \{1, 2, 3, 4, 5, 6\}$*

[1] De aquí en adelante se usará la pareja $[\Omega, p]$, teniendo en cuenta que Ω es finito y por esta razón induce de forma unívoca un espacio en el sentido de Kolmogorov.

El dado está cargado de la siguiente forma:

$$\begin{aligned} p(1) &= p(2) = p(3) = \frac{1}{4} \\ p(4) &= p(5) = p(6) = \frac{1}{12} \end{aligned}$$

cumpliendose que $p(1) + p(2) + ... + p(6) = 1$. Ahora se quiere obtener $P(A)$ en donde el evento $A =$ *la cara es par,* entonces $A = \{2, 4, 6\}$ por tanto:

$$\begin{aligned} P(A) &= \sum_{i \in A} p(i) \\ &= p(2) + p(4) + p(6) \end{aligned}$$

como el dado esta cargado entonces $P(A) = \frac{1}{4} + \frac{1}{12} + \frac{1}{12} = \frac{5}{12}$. Así ya no existe el problema de la definición clásica en donde se exigía que los eventos fueran equiprobables. Siguiendo la definición axiomática sólo se tiene que cumplir que $P(\Omega) = 1$.

Ejemplo 6 ***Lanzar un par de dados honestos.***

Sea

$$\Omega = \left\{ \begin{array}{c} (1,1), (1,2), (1,3), (1,4), (1,5)\ , (1,6,), \\ (2,1), (2,2), (2,3), (2,4), (2,5)\ , (2,6,), \\ (3,1), (3,2), (3,3), (3,4), (3,5)\ , (3,6,), \\ (4,1), (4,2), (4,3), (4,4), (4,5)\ , (4,6,), \\ (5,1), (5,2), (5,3), (5,4), (5,5)\ , (5,6,), \\ (6,1), (6,2), (6,3), (6,4), (6,5)\ , (6,6,) \end{array} \right\}$$

en donde (i, j) significa que en el primer dado se obtuvo la cara i y en el segundo se obtuvo la cara j, escrito de forma compacta

$$\Omega = \{(i, j)/i = 1, .., 6 \ y \ \ j = 1, .., 6\}.$$

$$\text{donde } p(i, j) = \frac{1}{36}$$

Sea W el evento "la suma de las caras es menor o igual que tres", es decir W esta formado por los siguientes elementos

$$W = \{(1, 1), (1, 2), (2, 1)\},$$

entonces

$$\begin{aligned} P(W) &= \sum_{(i,j) \in W} p(i, j) \\ &= p(1, 1) + p(1, 2) + p(2, 1) \\ &= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{3}{36} = \frac{1}{12} \end{aligned}$$

Sea X el evento "las caras de ambos dados son iguales", entonces

$$X = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$$

por lo tanto

$$P(X) = \sum_{(i,i) \in X} p(i, i) = \frac{6}{36} = \frac{1}{6}.$$

Los axiomas

Si Ω es finito cualquier subconjunto de Ω se conoce como *evento*, entonces se asigna a cada evento A un número $P(A)$ que se lee *"probabilidad del evento A"*. Está función $P : \Omega \rightarrow R$ es tal que debe satisfacer los siguientes tres axiomas:

Axioma 1 $P(A) \geq 0$

Axioma 2 $P(\Omega) = 1$

Axioma 3 *Si $A_1 \cap A_2 = \varnothing$, entonces $P(A_1 \cup A_2) = P(A_1) + P(A_2)$*

El tercer axioma se puede generalizar

Si $A_1, A_2, \ldots, A_n$ son mutuamente excluyentes (es decir que $A_i \cap A_j = \varnothing$ para todo $i \neq j$) entonces:

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

Estos axiomas fueron motivados por los enfoques de probabilidad tanto clásico como frecuentista.

Consecuencias elementales de los axiomas

Proposición 1 $P(A^c) = 1 - P(A)$

Demostración. Si $AUA^c = \Omega$ entonces de los axiomas 2 y 3 se tiene lo siguiente. Como $A \cap A^c = \varnothing$, entonces

$$1 = P(\Omega) = P(AUA^c) = P(A) + P(A^c).$$

■

Proposición 2 $P(\varnothing) = 0$

Demostración. Sabemos que $\varnothing \cup \Omega = \Omega$ y ambos son mutuamente excluyentes entonces

$$P(\varnothing \cup \Omega) = P(\varnothing) + P(\Omega).$$

■

Proposición 3 *Si $A_1, A_2, \ldots\ldots, A_n$ son eventos mutuamente excluyentes (es decir $A_i \cap A_j = \varnothing$ si $i \neq j$) entonces*

$$P(A_1 U A_2 \cup \ldots\ldots \cup A_n) = P(A_1) + P(A_2) + \ldots\ldots + P(A_n)$$

para cualquier conjunto A y B

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \leqq P(A) + P(B)$$

Demostración. Para probar esto basta con escribir $A \cup B$ y B como uniones de eventos mutuamente excluyentes, y en la figura 1-1, se representan dichos eventos.

$$\begin{aligned} A \cup B &= A \cup (A^c \cap B) \text{ ver figura 1-1 (a)} \\ P(A \cup B) &= P(A) + P(A^c \cap B) \\ P(A^c \cap B) &= P(A \cup B) - P(A), \end{aligned}$$

por otro lado

$$\begin{aligned} B &= (A \cap B) \cup (A^c \cap B) \ \text{ ver figura 1-1 (b)} \\ P(B) &= P(A \cap B) + P(A^c \cap B) \\ P(A^c \cap B) &= P(B) - P(A \cap B), \end{aligned}$$

de donde

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

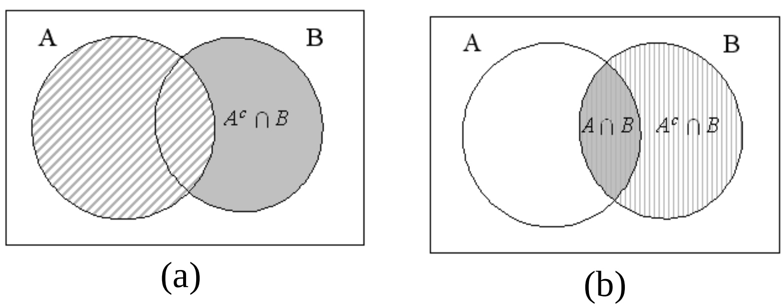


Figura 1-1

■

Proposición 4 *Si* $B \subset A$ *entonces*

$$P(A) = P(B) + P(A \cap B^c) \geqq P(B)$$

Demostración. Para probarlo basta escribir A como unión de eventos mutuamente excluyentes $A = B \cup (A \cap B^c)$ entonces

$$P(A) = P(B) + P(A \cap B^c) \text{ ver figura 1-2.}$$

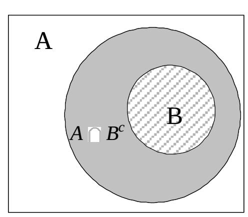


Figura 1-2

■

Probabilidad condicional

Definición 6 *Sean A y B dos eventos en el espacio $[\Omega, p]$. Cuando se desea determinar la probabilidad de un evento A dado que ya ocurrió un evento B, se le conoce como probabilidad condicional.*

La notación que se usará para la probabilidad condicional será:
$P(A|B)$ que se lee *"La probabilidad de A dado B"*.

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ \text{si } P(B) &> 0 \end{aligned}$$

análogamente se puede decir que si $P(A) > 0$ entonces

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

Ahora se demostrará que dado un evento fijo B, la probabilidad condicional es en sí una probabilidad. Para ello se verifica que satisface los axiomas de probabilidad.

1.- Demostrar que $P(A|B) \geq 0$.

Demostración. Esta condición obviamente se satisface pues

$$P(B) > 0 \text{ y } P(A \cap B) \geq 0.$$

■

2.- Demostrar que $P(\Omega|B) = 1$

Demostración. Como $B \subset \Omega$, entonces

$$P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

■

3.- Demostrar que si A_1 y A_2 son mutuamente excluyentes, entonces

$$P\left[(A_1 U A_2)|B\right] = P(A_1|B) + P(A_2|B)$$

Demostración. Se observa que $A_1 \cap A_2 = \oslash$, por lo tanto $(A_1 \cap B) \cap (A_2 \cap B) = \oslash$ ver figura 2.

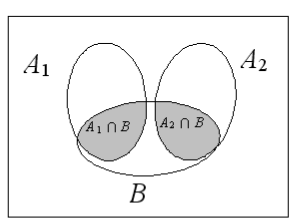


Figura 1-3

Entonces

$$\begin{aligned} P\left[(A_1 U A_2)|B\right] &= \frac{P\left[(A_1 \cup A_2) \cap B\right]}{P(B)} \\ &= \frac{P\left[(A_1 \cap B) U (A_2 \cap B)\right]}{P(B)} \\ &= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} \\ &= P(A_1|B) + P(A_2|B) \end{aligned}$$

Por lo tanto se sigue que todos los resultados que involucran probabilidades se cumplen también para probabilidades condicionales. ■

Propiedades

Proposición 5 *Demostrar que* $P(A|A) = 1$.

Demostración. $P(A|A) = \frac{P(A\cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$ ■

Proposición 6 *Si* $B \subset A$ *entonces* $P(A|B) = 1$.

Demostración. $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$ ■

Proposición 7 *Si* $A \subset B$ *entonces* $P(A|B) = \dfrac{P(A)}{P(B)}$.

Demostración. $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(A)}{P(B)} \geq P(A)$ ■

Ejemplos

Ejemplo 7 *En una urna se tienen 3 bolas blancas* b_1, b_2, b_3 *y 2 bolas rojas* r_1, r_2*. Si se extraen dos bolas. ¿Cuál es la probabilidad de obtener roja en la segunda extracción dado que la primera bola fue blanca? El conjunto* Ω *equiprobable es el siguiente:*

$$\Omega = \left\{ \begin{array}{cccc} (b_1, r_1) & (b_1, r_2) & (b_1, b_2) & (b_1, b_3) \\ (b_2, r_1) & (b_2, r_2) & (b_2, b_1) & (b_2, b_3) \\ (b_3, r_1) & (b_3, r_2) & (b_3, b_1) & (b_3, b_2) \\ (r_1, r_2) & (r_1, b_1) & (r_1, b_2) & (r_1, b_3) \\ (r_2, r_1) & (r_2, b_1) & (r_2, b_2) & (r_2, b_3) \end{array} \right\}$$

Los eventos A y B estan determinados como sigue:
$A =$ La segunda extracción es roja.

$$A = \left\{ \begin{array}{ccc} (b_1, r_1) & (b_1, r_2) & \\ (b_2, r_1) & (b_2, r_2) & \begin{array}{c}(r_1, r_2) \\ (r_2, r_1)\end{array} \\ (b_3, r_1) & (b_3, r_2) & \end{array} \right\} \quad \text{entonces} \quad P(A) = \frac{8}{20}$$

$B =$ La primera extracción es blanca.

$$B = \left\{ \begin{array}{cccc} (b_1, r_1) & (b_1, r_2) & (b_1, b_2) & (b_1, b_3) \\ (b_2, r_1) & (b_2, r_2) & (b_2, b_1) & (b_2, b_3) \\ (b_3, r_1) & (b_3, r_2) & (b_3, b_1) & (b_3, b_2) \end{array} \right\} \text{ entonces } P(B) = \frac{12}{20}$$

$A \cap B =$La primera extracción es blanca y la segunda es roja.

$$A \cap B = \left\{ \begin{array}{cc} (b_1, r_1) & (b_1, r_2) \\ (b_2, r_1) & (b_2, r_2) \\ (b_3, r_1) & (b_3, r_2) \end{array} \right\} \text{ entonces } P(A \cap B) = \frac{6}{20}$$

por lo tanto aplicando la fórmula de probabilidad condicional se tiene lo siguiente:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{6}{20}}{\frac{12}{20}} = \frac{6}{12} = \frac{1}{2}.$$

Otro camino para resolver este problema es restringir Ω a B

Ejemplo 8 *En un estudio de política internacional, se preguntó a una muestra de 2471 personas acerca de la atención que prestaban a las noticias, cuestiones económicas, o acontecimientos que suceden en otros países. En la siguiente tabla se clasifican las 2471 personas según su nivel de estudios (bajo, medio, alto) y el interés que muestran por los acontecimientos internacionales (ninguno, poco, bastante, mucho).*

Interés/Nivel de Estudios	Bajo (B)	Medio (M)	Alto (A)	Total
Mucho (X)	52	99	86	237
Bastante (Y)	368	664	233	1265
Poco (Z)	332	382	40	754
Ninguno (W)	138	66	11	215
Total	890	1211	370	2471

Se calculan las siguientes probabilidades:

Solución 1 *Probabilidad de que un individuo preste mucha o bastante atención si su nivel cultural es alto:*

$$P\left[(X \cup Y)\,|A\right] = \frac{P\left[(X \cup Y) \cap A\right]}{P(A)} = \frac{86 + 23}{370} = 0{,}862$$

Probabilidad de que un individuo preste mucha o bastante atención si su nivel cultural es medio:

$$P\left[(X \cup Y)\,|M\right] = \frac{P\left[(X \cup Y) \cap A\right]}{P(M)} = \frac{99 + 664}{1211} = 0{,}630$$

Probabilidad de que un individuo preste mucha o bastante atención si su nivel cultural es bajo:

$$P\left[(X \cup Y)\,|B\right] = \frac{P\left[(X \cup Y) \cap A\right]}{P(B)} = \frac{52 + 368}{890} = 0{,}471$$

Se observa que a medida que el nivel cultural es mayor, es mayor la probabilidad de que preste mucha o bastante interés en los acontecimientos internacionales.

Ejemplo 9 *En el experimento de lanzar un dado honesto. Se desea saber la probabilidad del evento la cara es dos dado que ya sucedió el evento la cara es par. Sea* $A = \{2\}$ *y* $P(A) = \frac{1}{6}$ *así como* $B = \{2, 4, 6\}$ *y* $P(B) = \frac{3}{6}$.

Como podemos observar $A \subset B$ por lo tanto utilizando la proposición 7:

$$P(A|B) = \frac{P(A)}{P(B)} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{6}{18} = \frac{1}{3}.$$

Ejemplo 10 *Refiriendonos a la radiografía ocupacional de 2004. Ver pag. 119 Se sabe que un individuo pertenece a la población no asalariada* (PNA) *¿Cuál es la probabilidad de que sea trabajador por cuenta propia* (TCP)*?*

Es claro que $(TCP) \subset (PNA)$ por lo tanto:

$$P(TCP|PNA) = \frac{P(TCP)}{P(PNA)} = \frac{10,361,154}{15,419,363} = 0{,}67$$

Por lo tanto el 67 % de la población no asalariada es trabajador por cuenta propia.

Teorema de multiplicación

Teorema 1 *Sea* $[\Omega, p]$ *un espacio de probabilidad, y sean*

$$A_1, A_2, A_3, ..., A_n$$

eventos de tal forma que

$$P(A_1 \cap A_2 \cap ... \cap A_r) > 0 \text{ con } 1 \leq r \leq n-1$$

entonces

$$\begin{aligned} P(A_1 \cap A_2 \cap ... \cap A_n) &= P(A_1)P(A_2|A_1)P\left(A_3|A_1 \cap A_2\right)... \\ &\quad ...P\left(A_n|A_1 \cap A_2 \cap ... \cap A_{n-1}\right) \end{aligned}$$

Demostración. La demostración será por inducción sobre n.

Si $n = 2$, sean A_1, A_2 eventos, en donde $P(A_1) > 0$ entonces la probabilidad condicional esta bien definida $P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}$ de donde

$$P(A_1 \cap A_2) = P(A_1)P(A_2|A_1).$$

Se supone que es cierto para $n = k$ con $k \geq 2$, entonces sean $A_1, A_2, ..., A_k$ eventos para los cuales $P(A_1 \cap A_2 \cap ... \cap A_{k-1}) > 0$; entonces

$$\begin{aligned} P(A_1 \cap A_2 \cap ... \cap A_k) &= P(A_1)P(A_2|A_1)P\left(A_3|A_1 \cap A_2\right)... \\ &\quad ...P\left(A_k|A_1 \cap A_2 \cap ... \cap A_{k-1}\right) \end{aligned}$$

Y se demuestra para $n = k + 1$

Sean $A_1, A_2, ..., A_k, A_{k+1}$ eventos tales que

$$P(A_1 \cap A_2 \cap ... \cap A_k) > 0$$

entonces

$$P\left[(A_1 \cap A_2 \cap ... \cap A_k) \cap A_{k+1}\right] =$$

$$\begin{aligned} &= P(A_1 \cap A_2 \cap ... \cap A_k) P\left(A_{k+1} | A_1 \cap A_2 \cap ... \cap A_k\right) \\ &= P(A_1) P(A_2 | A_1) P\left(A_3 | A_1 \cap A_2\right) ... \\ &\quad ... P\left(A_k | A_1 \cap A_2 \cap ... \cap A_{k-1}\right) P\left(A_{k+1} | A_1 \cap A_2 \cap ... \cap A_k\right) \end{aligned}$$

por lo que la fórmula general para el teorema de multiplicación ha quedado demostrada. ■

Teorema de probabilidad total

Teorema 2 *Sea* $[\Omega, p]$ *un espacio de probabilidad, si*

$$A_1, A_2, A_3, ..., A_n$$

es un conjunto de eventos mutuamente excluyentes que satisfacen

$$\Omega = A_1 \cup A_2 \cup A_3 \cup, ..., \cup A_n$$

y $P(A_j) > 0$ *para todo* $j = 1, ..., n$*; entonces para todo* X

$$P(X) = \sum_{i=1}^{n} P(A_i) P(X | A_i)$$

Demostración. Sea $X = X \cap \Omega$, entonces

$$X = X \cap (A_1 \cup A_2 \cup A_3 \cup, ..., \cup A_n)$$

y por la ley distributiva

$$X = (X \cap A_1) \cup (X \cap A_2) \cup ... \cup (X \cap A_n).$$

Por otro lado se sabe que $(X \cap A_i) \cap (X \cap A_j) = \varnothing$ son ajenos, es decir $A_i \cap A_j = \varnothing$ para todo $i \neq j$, ya que

$$\begin{aligned} (X \cap A_i) \cap (X \cap A_j) &= A_i \cap (X \cap (X \cap A_j)) \\ &= A_i \cap X \cap A_j \\ &= X \cap A_i \cap A_j \\ &= X \cap \varnothing = \varnothing \end{aligned}$$

por lo tanto

$$P(X) = P(X \cap A_1) + P(X \cap A_2) + ... + P(X \cap A_n)$$

y por la probabilidad condicional

$$P(X) = P(A_1) P(X | A_1) + P(A_2) P(X | A_2) + ... + P(A_n) P(X | A_n)$$

por lo que el teorema ha quedado demostrado. ■

Teorema de Bayes

Teorema 3 *Sea $[\Omega, p]$ un espacio de probabilidad, si $A_1, A_2, A_3, ..., A_n$ es un conjunto de eventos mutuamente excluyentes que satisfacen*

$$\Omega = A_1 \cup A_2 \cup A_3 \cup, ..., \cup A_n$$

y $P(A_j) > 0$ para todo $j = 1, ..., n$; entonces para todo X para el cual $P(X) > 0$ se cumple que

$$P(A_i|X) = \frac{P(A_i)P(X|A_i)}{\sum_{i=1}^{n} P(A_i)P(X|A_i)}.$$

Demostración. De acuerdo con la probabilidad condicional se tiene que

$$\begin{aligned} P(A_i \cap X) &= P(A_i)P(X|A_i) \\ &= P(X)P(A_i|X) \end{aligned}$$

de donde $P(A_i|X) = \frac{P(A_i)P(X|A_i)}{P(X)}$ y por el teorema de probabilidad total

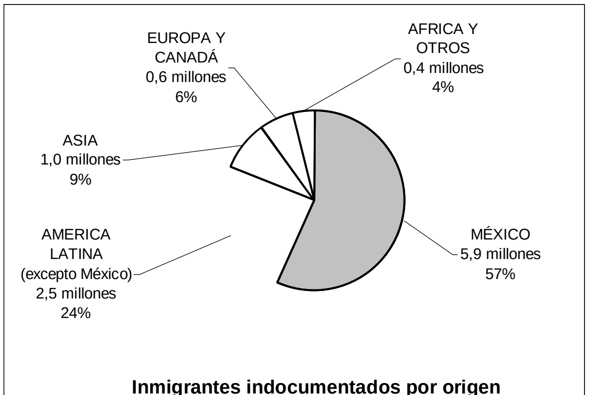
$$P(A_i|X) = \frac{P(A_i)P(X|A_i)}{P(A_1)P(X|A_1) + P(A_2)P(X|A_2) + ... + P(A_n)P(X|A_n)}$$

■

El siguiente ejemplo aplicado a datos reales involucra el teorema de multiplicación y el teorema de Bayes.

Ejemplo 11 ***Inmigrantes Indocumentados en Estados Unidos***

En Estados Unidos el Pew Hispanic Center dió a conocer en el mes de marzo de 2005 los principales resultados de su estudio ESCUP por sus siglas en inglés (estimaciones del tamaño y características de la población indocumentada). En la categoría indocumentados según su origen se tiene que hasta marzo del 2004 había 10.3 millones de indocumentados.[2]



Inmigrantes indocumentados por origen

[2]Fuente UNITE Unidad Técnica de Economía S.A. de C.V. unite@i.com.mx

Se sabe que el 1 % de la población de indocumentados de Europa y Canadá, el 2 % de África, el 3 % de Asia, el 8 % de América Latina y por último el 18 % de México alguna vez ha sufrido algún tipo de discriminación racial. *¿Cuál es la probabilidad de que un individuo de la localidad haya sufrido algún tipo de discriminación racial?.* Si se denota a los eventos

EYC = *pertenecer a la población de indocumentados de Europa y Canadá*

Solución 2 AF = *pertenecer a la población de indocumentados de África*

AS = *pertenecer a la población de indocumentados de Asia*

AL = *pertenecer a la población de indocumentados de América Latina*

M = *pertenecer a la población de indocumentados de México*

D = *sufrir algún tipo de discriminación*

$$\begin{aligned}
P(D) &= P(D|EYC)P(EYC) + P(D|AF)P(AF) + P(D|AS)P(AS) + \\
&\quad + P(D|AL)P(AL) + P(D|M)P(M) \\
&= [0{,}01 * 0{,}06] + [0{,}02 * 0{,}04] + [0{,}03 * 0{,}09] + [0{,}08 * 0{,}27] + [0{,}18 * 0{,}57] \\
&= 0{,}1283
\end{aligned}$$

es decir un 12.8 % de la población total de indocumentados ha sufrido algún tipo de discriminación.

Ejemplo 12 *Utilizando la misma información del ejemplo anterior, si sabemos que un individuo ha sufrido algún tipo de maltrato ¿Cuál es la probabilidad de que sea de México?*

$$\begin{aligned}
P(M|D) &= \frac{P(M \cap D)}{P(D)} \\
&= \frac{P(D|A)P(A)}{P(D)} \\
&= \frac{[0{,}18 * 0{,}57]}{0{,}1283} = 0{,}7996
\end{aligned}$$

Es decir el 79.9 % de la población indocumentada de México en Estados Unidos ha sufrido algún tipo de discriminación racial.

Independencia entre eventos

El concepto de independencia ocupa una posición central en la teoría de la probabilidad, ya que justifica el desarrollo matemático de esta teoría pero no como un tópico de teoría de la medida sino como una disciplina separada. Lo anterior será apreciado más tarde, por ahora sólo se discutiran algunas de sus propiedades.

Intuitivamente se dice que A es independiente de B si la probabilidad de A no depende de la ocurrencia de B. Por ejemplo al lanzar dos dados de diferentes colores, lo que salga en el dado rojo no depende de lo que salga en el dado azul. La definición formal es la siguiente:

Proposición 8 *Sea $[\Omega, p]$ un espacio de probabilidad y sean A y B eventos, se dice que A y B son independientes si y sólo si:*

$$P(A \cap B) = P(A)P(B).$$

Observación 4 *El concepto de independencia se puede ver desde el punto de vista de la probabilidad condicional de la siguiente forma:*

$$\begin{aligned} P(A|B) &= P(A) \text{ si } P(B) > 0 \\ P(B|A) &= P(B) \text{ si } P(A) > 0 \end{aligned}$$

pues ambas implican que $P(A \cap B) = P(A)P(B)$

Observación 5 *Independencia y mutuamente excluyentes son 2 propiedades distintas pero están relacionadas.*

- *Dos eventos mutuamente excluyentes son independientes si y sólo si*

  $$P(A)P(B) = 0$$

  es decir que alguna de las dos probabilidades tiene que ser cero.

- *Si se sabe que $P(A) \neq 0$ y $P(B) \neq 0$ entonces:*

  A y B son independientes $\Rightarrow$ A y B no son mutuamente excluyentes

  A y B son mutuamente excluyentes $\Rightarrow$ A y B no son independientes

Ejemplo 13 *Se considera el lanzamiento de dos dados. Y los eventos estan definidos de la siguiente forma:*
$A =$La suma de las caras de ambos dados es impar $P(A) = \frac{1}{2}$
$B =$Uno en la primera entrada $P(B) = \frac{1}{6}$
$C =$La suma de ambas caras es siete $P(C) = \frac{1}{6}$.
Encontrar las siguientes probabilidades condicionales $P(A|B)$, $P(A|C)$ y $P(C|B)$

$$\begin{aligned} P(A|B) &= \frac{1}{2} = P(A) \text{ por lo tanto } A \text{ y } B \text{ son independientes} \\ P(A|C) &= 1 \neq P(A) = \frac{1}{2} \text{por lo tanto } A \text{ y } C \text{ no son independientes} \\ P(C|B) &= \frac{1}{6} = P(C) \text{ por lo tanto } B \text{ y } C \text{ son independientes.} \end{aligned}$$

Teorema 6 *Si A y B son dos eventos independientes en un espacio de probabilidad $[\Omega, p]$, entonces:*

$$\begin{aligned} P(A \cap B^c) &= P(A)\, P(B^c) \\ P(A^c \cap B) &= P(A^c)\, P(B) \\ P(A^c \cap B^c) &= P(A^c)\, P(B^c) \end{aligned}$$

Demostración.

$$\begin{aligned}
P(A \cap B^c) &= P(A) - P(A \cap B) \\
&= P(A) - P(A)P(B) \\
&= P(A)[1 - P(B)] \\
&= P(A)P(B^c) \text{ ver figura 2}
\end{aligned}$$

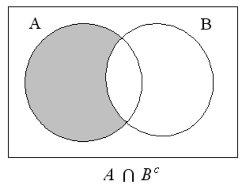


Figura 1-4

$$\begin{aligned}
P(A^c \cap B) &= P(B) - P(A \cap B) \\
&= P(B) - P(A)P(B) \\
&= P(B)[1 - P(A)] \\
&= P(A^c)P(B) \text{ ver figura 2.}
\end{aligned}$$

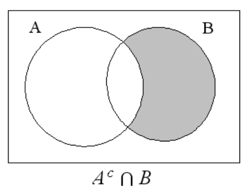


Figura 1-5

Por último

$$\begin{aligned}
P(A^c \cap B^c) &= P(A^c|B^c)P(B^c) \\
&= P(A^c)P(B^c).
\end{aligned}$$

■

Definición 7 ***Independencia para más de dos eventos.*** *Se considera* $[\Omega, p]$ *un espacio de probabilidad, y sean* $A_1, A_2, A_3, ..., A_n$ *eventos. Se dice que esos eventos serán independientes si y sólo si*

$$\begin{aligned}
P(A_i \cap A_j) &= P(A_i)P(A_j) \quad \textit{para } i \neq j \\
\\
P(A_i \cap A_j \cap A_k) &= P(A_i)P(A_j)P(A_k) \quad \textit{para } i \neq j,\ j \neq k, i \neq k \\
&\vdots \\
P\left[\cap_{i=1}^{n} A_i\right] &= \Pi_{i=1}^{n} P(A_i)
\end{aligned}$$

Observación 7

$$\begin{aligned} \textit{Independencia total} &\Rightarrow \textit{independencia por pares} \\ \textit{Independencia por pares} &\nRightarrow \textit{independencia total.} \end{aligned}$$

Ejemplo 14 ***Independencia por pares no implica independencia.*** *Se considera el lanzamiento de dos dados. Y los eventos estan definidos de la siguiente forma:*

$$\begin{aligned} A_1 &= \textit{La cara es impar en el primer dado } P(A_1) = \frac{1}{2} \\ A_2 &= \textit{La cara es impar en el segundo dado } P(A_2) = \frac{1}{2} \\ A_3 &= \textit{La suma de ambas caras da impar } P(A_3) = \frac{1}{2} \end{aligned}$$

$P(A_1|A_2) = \frac{1}{2} = P(A_1)$ entonces A_1 y A_2 son independientes
$P(A_1|A_3) = \frac{1}{2} = P(A_1)$ entonces A_1 y A_3 son independientes
$P(A_3|A_2) = \frac{1}{2} = P(A_3)$ entonces A_3 y A_2 son independientes.

Solución 3 *Pero los tres no son independientes pues*

$$P(A_1 \cap A_2 \cap A_3) \neq P(A_1)P(A_2)P(A_3)$$

ya que

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \\ &= \frac{1}{2} * \frac{1}{2} * 0 = 0 \end{aligned}$$

y

$$P(A_1)P(A_2)P(A_3) = \frac{1}{8}.$$

En resumen, se puede concluír que no es fácil hacer un cálculo directo para obtener $P(A \cap B)$,entonces se usaran los conceptos de independencia y probabilidad condicional pues sirven para facilitar este cálculo. Cuando se sabe que existe independencia, entonces $P(A \cap B) = P(A)P(B)$, y si no existe independencia $P(A \cap B) = P(A|B)P(B)$, y encontrar estas probabilidades $P(A)$, $P(B)$, $P(A|B)$ es relativamente más fácil.

Espacios producto de probabilidad

A continuación se definen los espacios producto de probabilidad, los cuales serán de utilidad ya que son una forma sencilla de construír gran variedad de eventos independientes.

Definición 8 *Sean $E_1 = [\Omega_1, p_1]$ y $E_2 = [\Omega_2, p_2]$ dos espacios de probabilidad finitos, es decir Ω_1 es el espacio muestra que tiene asociada la función $p_1 : \Omega_1 \longrightarrow R_+$ tal que $\sum_{i=1}^{n} p(a_i) = 1$, así mismo el espacio muestra Ω_2 tiene asociada la función $p_2 : \Omega_2 \longrightarrow R_+$ tal que $\sum_{j=1}^{m} p(a_j) = 1$. Entonces un espacio producto de probabilidad $E = E_1 \times E_2$, es un sistema $E = [\Omega, P]$, con $\Omega = \Omega_1 \times \Omega_2$, y donde se define a P como una función $P : \Omega \longrightarrow R_+$ tal que $P(a_i, a_j) = p_1(a_i)p_2(a_j)$ con $a_i \in \Omega_1$ y $a_j \in \Omega_2$.*

Es fácil demostrar que $\sum_{i=1}^{n} \sum_{j=1}^{m} p_1(a_i)p_2(a_j) = 1$.

Observación 8 *Sean $A_1 \subset \Omega_1$ y $A_2 \subset \Omega_2$, se consideran ahora los productos cartesianos generados por $A_1 \times \Omega_2$ y $\Omega_1 \times A_2$, ambos eventos son subconjuntos de Ω, y son independientes en E.*

Ejemplo 15 *El lanzamiento de un dado y de una moneda, generan el siguiente espacio producto $E = [\Omega, P]$. En donde $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$, y tal que $p_1(i) = \frac{1}{6}$ con $i = 1, ..., 6$ y sea $\Omega_2 = \{a, s\}$, tal que $p_2(a) = \frac{2}{3}$, y $p_2(s) = \frac{1}{3}$.*

$$\Omega = \Omega_1 \times \Omega_2 = \left\{ \begin{array}{c} (1,a), (2,a), (3,a), (4,a), (5,a), (6,a), \\ (1,s), (2,s), (3,s), (4,s), (5,s), (6,s) \end{array} \right\},$$

en donde las probabilidades de cada uno de los elementos del espacio producto están dadas por:

$$\begin{array}{ll} P(1,a) = p_1(1)p_2(a) = \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9} & P(1,s) = p_1(1)p_2(s) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18} \\ P(2,a) = p_1(2)p_2(a) = \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9} & P(2,s) = p_1(2)p_2(s) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18} \\ P(3,a) = p_1(3)p_2(a) = \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9} & P(3,s) = p_1(3)p_2(s) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18} \\ P(4,a) = p_1(4)p_2(a) = \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9} & P(4,s) = p_1(4)p_2(s) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18} \\ P(5,a) = p_1(5)p_2(a) = \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9} & P(5,s) = p_1(5)p_2(s) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18} \\ P(6,a) = p_1(6)p_2(a) = \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9} & P(6,s) = p_1(6)p_2(s) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18} \end{array},$$

es claro que al sumar estas probabilidades nos da uno.

Haciendo hincapié en la observación 8 y utilizando el ejemplo anterior, se ve cómo se generan los eventos independientes. Sea A_1 el evento al lanzar un dado se obtiene un número par, es decir $A_1 = \{2, 4, 6\}$, donde $A_1 \subset \Omega_1$ y A_2 el evento al lanzar una moneda cae águila, es decir $A_2 = \{a\}$, donde $A_2 \subset \Omega_2$. A continuación se detallan los elementos de $A_1 \times \Omega_2$ y $\Omega_1 \times A_2$, se puede ver claramente que ambos productos cartesianos son subconjuntos de Ω.

$$\begin{aligned} A_1 \times \Omega_2 &= \{(2,a), (4,a), (6,a), (2,s), (4,s), (6,s)\} \\ \Omega_1 \times A_2 &= \{(1,a), (2,a), (3,a), (4,a), (5,a), (6,a)\} \end{aligned}$$

A_1 y A_2 son independientes en E por construcción, pero lo podemos comprobar obteniendo la probabilidad de la intersección:

$$\begin{aligned} P(A_1 \cap A_2) &= \frac{3}{9} = \frac{1}{3} \\ P(A_1) * P(A_2) &= \frac{3}{6} * \frac{2}{3} = \frac{1}{3} \end{aligned}$$

Generalizando lo anterior podemos definir lo siguiente:

Definición 9 *Sean* $E_1 = [\Omega_1, p_1]$, $E_2 = [\Omega_2, p_2]\,, ..., E_n = [\Omega_n, p_n]\,,$ *n espacios de probabilidad finitos. El espacio producto de probabilidad de todos ellos, es un sistema* $E = [\Omega, P]$ *con* $\Omega = \Omega_1 \times \Omega_2 \times ... \times \Omega_n$*, en donde se define a P como una función* $P : \Omega \longrightarrow R_+$ *tal que* $P(a_1, a_2, ..., a_n) = p_1(a_1)p_2(a_2)...p_n(a_n)$.

También aquí se puede demostrar que

$$\sum_{(a_1,a_2,...,a_n)\in\Omega} P(a_1, a_2, ..., a_n) = 1.$$

Observación 9 *Sean* $A_1 \subset \Omega_1$, $A_2 \subset \Omega_2$,...,$A_n \subset \Omega_n$*, considérense los siguientes productos cartesianos*

$$\begin{aligned}
\overline{A_1} &= A_1 \times \Omega_2 \times ... \times \Omega_n \subset \Omega_1 \times \Omega_2 \times ... \times \Omega_n \\
\overline{A_2} &= \Omega_1 \times A_2 \times ... \times \Omega_n \subset \Omega_1 \times \Omega_2 \times ... \times \Omega_n \\
&\vdots \\
\overline{A_n} &= \Omega_1 \times \Omega_2 \times ... \times A_n \subset \Omega_1 \times \Omega_2 \times ... \times \Omega_n
\end{aligned}$$

todos los $\overline{A_i}$ *también son eventos independientes en* E.

Capítulo 3

Teorema de DeMoivre-Laplace

Este capítulo está dedicado básicamente a la demostración detallada del teorema de De Moivre Laplace, dicho teorema viene motivado por los experimentos Bernoulli. Por esta razón, se empieza el capítulo considerando cómo son los ensayos o experimentos Bernoulli, después se enuncia el teorema de DeMoivre-Laplace, es preciso comentar que se tuvo especial cuidado en la demostración de este teorema, en la búsqueda de que conceptos como puede ser en este caso, el de convergencia uniforme sea captado del mejor modo posible por el estudiante, se incluyen algunos tipos de aplicaciones al teorema para entender su significado. También se demuestra el teorema integral de DeMoivre-Laplace y por último, se enuncia un resultado relativo a la ley de los grandes números.

Los experimentos considerados por Jacob Bernoulli

Es bien sabido del análisis combinatorio que si un conjunto tiene "n" elementos, entonces el número de subconjuntos con m elementos es igual a el coeficiente binomial

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} \tag{3.1}$$

Por ejemplo si queremos saber cuantos son todos los subconjuntos que se pueden formar de 2 elementos a partir del conjunto $\{a, b, c, d\}$, sabemos que serían los siguientes seis subconjuntos

$$\{a, b\}, \quad \{a, c\}, \quad \{a, d\}, \quad \{b, c\}, \quad \{b, d\}, \quad \{c, d\},$$

pero generalmente es difícil enumerarlos sobretodo en conjuntos con más elementos. Para ello se usa la fórmula 3.1, así en nuestro ejemplo tenemos $n = 4$ y $m = 2$ entonces $\binom{4}{2} = \frac{4!}{2!2!} = 6.$

El resultado anterior será usado para encontrar la probabilidad de que un evento ocurra m veces en n sucesos independientes de un experimento con espacio muestra Ω. Este problema es esencialmente el mismo que el problema de obtener m aguilas al tirar n veces una moneda.

Se tiene un experimento con espacio muestra Ω y un evento A que ocurre con probabilidad p, en tanto que A^c sucede con probabilidad $q = 1 - p$, es decir

$$P(A) = p \text{ y } P(A^c) = q \text{ donde } p + q = 1.$$

Repetimos el experimento n veces resultando un espacio producto que se denota con Ω^n,

$$\Omega^n = \Omega \times ... \times \Omega$$

quisieramos determinar la probabilidad $P_n(m)$ que corresponde a:

$$P_n(m) = P(A \text{ ocurra } m \text{ veces en cualquier orden en } n \text{ lanzamientos}).$$

Después de realizar n veces el experimento básico, se puede considerar a la secuencia particular en la que A se llevó acabo m veces, seguida de $n - m$ ocurrencias de A^c, así mismo está secuencia corresponde a un elemento $n-$dimensional en el espacio muestra Ω^n.

Puesto que se supone la independencia en los lanzamientos, la probabilidad de una secuencia como la mencionada es:

$$P(A)P(A)...P(A)P(A^c)P(A^c)...P(A^c) = p^m q^{n-m}, \tag{3.2}$$

en donde $P(A)$ se multiplica m veces y $P(A^c)$ se multiplica $n - m$ veces. Evidentemente existen otras secuencias particulares en las que se producirá m veces A y $n - m$ veces el suceso A^c, pero la probabilidad de cada suceso viene dada por la fórmula 3.2, por lo tanto la probabilidad del total del secuencias corresponde al coeficiente binomial multiplicado por $p^m q^{n-m}$, es decir:

$$P_n(m) = \binom{n}{m} p^m q^{n-m}, \tag{3.3}$$

y sustituyendo el coeficiente binomial, también podemos ver la fórmula como:

$$P_n(m) = \frac{n!}{m!(n-m)!} p^m q^{n-m}.$$

Ejemplo 16 *Se lanza una moneda honesta* 3 *veces y se desea calcular la probabilidad de que el evento "águila" se obtenga* 2 *veces. El espacio de probabilidad del lanzamiento de una moneda es* $\Omega = \{\acute{a}guila, sol\}$ *con* $p(\acute{a}guila) = \frac{1}{2}$ *y* $p(sol) = \frac{1}{2}$. *Ahora el espacio muestra es:*

$$\begin{aligned} \Omega^3 &= \Omega \times \Omega \times \Omega \\ &= \{(\acute{a}guila, \acute{a}guila, \acute{a}guila), (\acute{a}guila, \acute{a}guila, sol), \\ &\quad (\acute{a}guila, sol, \acute{a}guila), (\acute{a}guila, sol, sol), (sol, \acute{a}guila, \acute{a}guila), \\ &\quad (sol, \acute{a}guila, sol), (sol, sol, \acute{a}guila), (sol, sol, sol)\} \end{aligned}$$

Por lo tanto $P(\textit{El evento águila se obtenga 2 veces}) = \frac{3}{8}$. *Frecuentemente los espacios de probabilidad que se generan son mucho mayores por lo tanto no es conveniente enumerar cada elemento* *del espacio muestra, es por ello que se usa la fórmula 3.2. Si se lanza la moneda una sola vez el evento $A = \{águila\}$ ocurre con probabilidad $\frac{1}{2}$. Por lo tanto*

$$P(A) = \frac{1}{2}, \quad P(A^c) = \frac{1}{2}, \quad n = 3 \ y \ m = 2$$

entonces

$$p_3(2) = \left(\begin{array}{c} 3 \\ 2 \end{array} \right) \left(\frac{1}{2} \right)^2 \left(\frac{1}{2} \right)^1 = \frac{3}{8}.$$

Ejemplo 17 *Se lanza un dado honesto 5 veces. Y se desea calcular la probabilidad de que el evento "seis"se obtnga 2 veces.Si se lanza el dado una sola vez el evento $A = \{seis\}$ y ocurre con probabilidad $\frac{1}{6}$. Por lo tanto ,*

$$P(A) = \frac{1}{6}, \ \ P(A^c) = \frac{5}{6}, \ n = 5 \ y \ m = 2$$

entonces

$$p_5(2) = \left(\begin{array}{c} 5 \\ 2 \end{array} \right) \left(\frac{1}{6} \right)^2 \left(\frac{5}{6} \right)^3 = \frac{3}{8}.$$

Ejemplo 18 *La probabilidad de que algún artículo manufacturadoa sea defectuoso es* 0,005*. ¿Cuál es la probabilidad de que de* 10,000 *artículos elegidos al azar exactamente* 40 *sean defectuosos y no más de* 70 *sean defectuosos? Sea el evento $A = 40$ son defectuosos, $n = 10,000 \quad m = 40$ y $p = 0{,}005$, por lo tanto*

$$P(A) = \left(\begin{array}{r} 10000 \\ 40 \end{array} \right) (0{,}005)^{40} (0{,}995)^{9960}.$$

Sea el evento $B =$ No más de 70 *sean defectuosos, entonces obtenemos las siguientes probabilidades*

$$\begin{aligned}
P(0 \ defectuosos) &= \left(\begin{array}{r} 10000 \\ 0 \end{array} \right) (0{,}005)^{0} (0{,}995)^{10000-0} \\
P(1 \ defectuosos) &= \left(\begin{array}{r} 10000 \\ 1 \end{array} \right) (0{,}005)^{1} (0{,}995)^{10000-1} \\
&\vdots \\
P(70 \ defectuosos) &= \left(\begin{array}{r} 10000 \\ 70 \end{array} \right) (0{,}005)^{70} (0{,}995)^{10000-70}
\end{aligned}$$

así la probabilidad del evento B es

$$\begin{aligned}
P(B) &= \sum_{m=0}^{70} P_n(m) \\
&= \sum_{m=0}^{70} \left(\begin{array}{r} 10000 \\ m \end{array} \right) (0{,}005)^{m} (0{,}995)^{10000-m}.
\end{aligned}$$

En este último ejemplo se puede ver que un cálculo directo de la probabilidad con la fórmula 3.2, presenta dificultades cuando n es grande, trataremos de resolver este conflicto más adelante en la sección del teorema Local de De Moivre Laplace.

Teorema de De Moivre-Laplace

El problema que queremos resolver es calcular la probabilidad de que ocurra el evento éxito al realizar el experimento de Bernoulli n veces, es decir calcular $P_n(m)$.

En la discusión con aplicaciones numéricas de las secciones anteriores se llega a la conclusión de que para valores grandes de n, el cálculo de probabilidades $P_n(m)$ usando la fórmula

$$P_n(m) = \frac{n!}{m!(n-m)!} p^m q^{n-m},$$

presenta dificultades considerables, pues involucra el cálculo de factoriales de números grandes.

Es necesario entonces, determinar estas probabilidades con el suficiente grado de exactitud, sin tener que calcular dichos factoriales. Un resultado de este estilo fué desarrollado por De Moivre en su libro "The Doctrine of Chances"(1733), para el caso especial en el cual $p = q = \frac{1}{2}$ y después generalizado por Laplace, en el cual p es un valor arbitrario entre cero y uno y $q = 1 - p$. Es por esto que a De Moivre se le relaciona con el problema de la moneda honesta y a Laplace con el de la moneda cargada. Lo anterior se resume en el importante teorema de De Moivre-Laplace.

Para demostrar el teorema usaremos la conocida fórmula de Stirling,

$$s! = \sqrt{2\pi s} \cdot s^s e^{-s} e^{\theta_s},$$

en donde el exponente θ_s satisface la desigualdad $|\theta_s| \leqq \frac{1}{12s}$.

Con esta breve introducción se tienen las condiciones para enunciar el teorema, y la siguiente sección será dedicada a su demostración. Es importante comentar, que en la búsqueda de que quede muy explícito el significado del teorema, la demostración será muy detallada.

El teorema y su demostración

Antes de enunciar el teorema se establecerá una convención de notación. Sea $0 < p < 1$ y $q = 1 - p$, entonces si $n \geqq 1$ y m son dos naturales tales que $0 \leqq m \leqq n$, denotaremos a

$$x_{(n,m)} = \frac{m - np}{\sqrt{npq}},$$

en algunos casos para simplificar la notación se escribe x, en vez de $x_{(n,m)}$.

Teorema 10 (De Moivre-Laplace) *Sea $0 < p < 1$ y $q = 1 - p$, y sean $a < b$ dos reales. Entonces para todo $\varepsilon > 0$ existe un natural N_ε tal que si $n > N_\varepsilon$ se tiene que, cuando $0 \leqq m \leqq n$ son tales que $x_{(n,m)} \in [a, b]$, entonces*

$$\left| \frac{\sqrt{npq} P_n(m)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}} - 1 \right| < \varepsilon. \tag{3.4}$$

Demostración. *La demostración estará dividida en varios pasos. A continuación se presenta un esquema general de la estratégia a seguir. En el primer paso se muestra que la prueba de la* *desigualdad 3.4, se reduce a la prueba de la siguente desigualdad*

$$\left|\log \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}}\right| < \eta,$$

con η *convenientemente elegida. En los pasos 2 y 3, buscamos ver a qué equivale la expresión*

$$\log \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}}. \tag{3.5}$$

Del paso 4 hasta el paso 9, estarán dedicados a acotar la expresión 3.5. Y por último, en el paso 10 se concluye la demostración.

Paso 1 *Sea* $\varepsilon > 0$, *construyamos* $\eta = \log(1+\varepsilon) > 0$. *Si*

$$\left|\log \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}}\right| < \eta \text{ entonces } \left|\frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}} - 1\right| < \varepsilon.$$

Demostración *Sea* $y = \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}} > 0$. *Supongamos que* $|\log\ y| < \eta$ *entonces*

$$-\eta < \log\ y < \eta,$$

como la exponencial es monótona creciente, lo anterior se puede expresar como

$$\begin{aligned} e^{-\eta} &< y < e^{\eta} \\ \frac{1}{e^{\eta}} &< y < e^{\eta}, \end{aligned}$$

haciendo $\eta = \log(1+\varepsilon)$, *se tiene que*

$$\frac{1}{1+\varepsilon} < y < 1+\varepsilon.$$

Podemos observar que $\frac{1}{1+\varepsilon} - (1-\varepsilon) = \frac{\varepsilon^2}{1+\varepsilon} > 0$, *de donde se puede decir que*

$$1-\varepsilon < \frac{1}{1+\varepsilon},$$

así pues

$$1-\varepsilon < y < 1+\varepsilon$$

de tal forma que

$$-\varepsilon < y-1 < \varepsilon$$

por lo tanto

$$|y-1| < \varepsilon.$$

Paso 2 *Si* $1 \leqq m < n$, *entonces*

$$P_n(m) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} \left(\frac{np}{m}\right)^m \left(\frac{nq}{n-m}\right)^{n-m} e^{\theta},$$

donde $|\theta| \leqq \frac{1}{12}\left(\frac{1}{n} + \frac{1}{m} + \frac{1}{n-m}\right)$.

Demostración *Utilizando la fórmula de Stirling se tiene que*

$$\begin{aligned}
P_n(m) &= \frac{n!}{m!(n-m)!} p^m q^{n-m} \\
&= \frac{\sqrt{2\pi n} n^n e^{-n} e^{\theta_n}}{\sqrt{2\pi m} m^m e^{-m} e^{\theta_m} \sqrt{2\pi(n-m)} (n-m)^{(n-m)} e^{-(n-m)} e^{\theta_{(n-m)}}} p^m q^{n-m} \\
&= \frac{\sqrt{2\pi n} n^n e^{-n} e^{\theta_n} e^{-\theta_m} e^{-\theta_{(n-m)}}}{\sqrt{2\pi m} m^m e^{-m} \sqrt{2\pi(n-m)} (n-m)^{(n-m)} e^{-(n-m)}} p^m q^{n-m} \\
&= \frac{\sqrt{n} \cdot n^n e^{-n} e^{\theta_n - \theta_m - \theta_{(n-m)}}}{\sqrt{2\pi m(n-m)} m^m e^{-m} (n-m)^{(n-m)} e^{-(n-m)}} p^m q^{n-m},
\end{aligned} \tag{3.6}$$

donde $|\theta_n| \leqq \frac{1}{n}$, $|\theta_m| \leqq \frac{1}{m}$, $\left|\theta_{(n-m)}\right| \leqq \frac{1}{n-m}$.

Al simplificar la expresión 3.6, y haciendo $\theta = \theta_n - \theta_m - \theta_{(n-m)}$, *se tiene*

$$P_n(m) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} \left(\frac{p}{m}\right)^m \left(\frac{q}{n-m}\right)^{n-m} n^n e^{\theta},$$

en donde

$$\begin{aligned}
|\theta| &= \left|\theta_n - \theta_m - \theta_{(n-m)}\right| \\
|\theta| &\leqq |\theta_n| + |\theta_m| + \left|\theta_{(n-m)}\right|
\end{aligned}$$

y por la fórmula de Stirling

$$|\theta| \leqq \frac{1}{12}\left(\frac{1}{n} + \frac{1}{m} + \frac{1}{n-m}\right).$$

Por otro lado, como $n^n = n^m n^{(n-m)}$, *entonces podemos escribir*

$$P_n(m) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} \left(\frac{np}{m}\right)^m \left(\frac{nq}{n-m}\right)^{n-m} e^{\theta},$$

donde

$$|\theta| \leqq \frac{1}{12}\left(\frac{1}{n} + \frac{1}{m} + \frac{1}{n-m}\right).$$

Así, el segundo paso ha quedado demostrado.

Paso 3 *Si $1 \leqq m < n$, entonces*

$$\log \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}} = S + R + \theta$$

donde

$$S = \log \sqrt{\frac{n}{m(n-m)}} + \log \sqrt{npq},$$

$$R = \log \left[\left(\frac{np}{m} \right)^m \left(\frac{nq}{n-m} \right)^{n-m} \right] + \frac{x^2}{2}$$

y

$$|\theta| \leqq \frac{1}{12} \left(\frac{1}{n} + \frac{1}{m} + \frac{1}{n-m} \right).$$

Demostración *Consideremos la expresión* $\log \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}}$, *usando las propiedades de los logaritmos, se deduce que:*

$$\begin{aligned}
\log \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}} &= \log P_n(m) - \log \left[\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}} \right] \\
&= \log P_n(m) - \log \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) + \log \sqrt{npq} \\
&= \log P_n(m) - \log \left(\frac{1}{\sqrt{2\pi}} \right) - \log e^{-\frac{x^2}{2}} + \log \sqrt{npq} \\
&= \log P_n(m) - \log \left(\frac{1}{\sqrt{2\pi}} \right) - \frac{x^2}{2} + \log \sqrt{npq}, \qquad (3.7)
\end{aligned}$$

recordemos que por el paso 2 se establece que

$$P_n(m) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} \left(\frac{np}{m} \right)^m \left(\frac{nq}{n-m} \right)^{n-m} e^{\theta},$$

así que el logaritmo de la expresión anterior es

$$\log P_n(m) = \log \frac{1}{\sqrt{2\pi}} + \log \sqrt{\frac{n}{m(n-m)}} + \log \left[\left(\frac{np}{m} \right)^m \left(\frac{nq}{n-m} \right)^{n-m} \right] + \theta, \qquad (3.8)$$

así pues sustituyendo 3.8 en 3.7, y simplificando se tiene que

$$\begin{aligned}
\log \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}} &= \log \sqrt{\frac{n}{m(n-m)}} + \log \sqrt{npq} + \log \left[\left(\frac{np}{m} \right)^m \left(\frac{nq}{n-m} \right)^{n-m} \right] + \frac{x^2}{2} + \theta \\
&= S + R + \theta
\end{aligned}$$

Por lo tanto, ha quedado demostrado.

Como ya se había señalado, ahora se planteamos el objetivo de acotar la expresión 3.5. Los pasos 4, 5 y 6, serán para asegurar que algunos términos no se hacen cero y otros detalles técnicos para evitar futuras complicaciones y así preparar los pasos finales. Los pasos 7, 8 y 9, estarán dedicados a acotar los términos S, R y θ, respectivamente.

Paso 4 *Sea N_0 un número natural tal que $N_0 > \text{máx}(4\alpha^2\frac{p}{q}, 4\alpha^2\frac{q}{p})$ donde $\alpha = \text{máx}(b, -a)$. Si $n > N_0$ y m son naturales tales que $0 \leqq m \leqq n$ y si $x_{(n,m)} \in [a, b]$ entonces*

$$\left|x_{(n,m)}\right| \sqrt{\frac{q}{np}} < \frac{1}{2} \text{ y } \left|x_{(n,m)}\right| \sqrt{\frac{p}{nq}} < \frac{1}{2}.$$

Consecuentemente,

$$1 + x_{(n,m)}\sqrt{\frac{q}{np}} > 0 \text{ y } 1 - x_{(n,m)}\sqrt{\frac{p}{nq}} > 0.$$

Demostración *Es fácil ver que α siempre será positivo, pues si $a > 0$ y $a < b$, se deduce que $b > 0$; por el contrario si $a < 0$ entonces $-a > 0$, por lo que $\alpha > 0$.*

Observación 11 *Si $a \leqq x \leqq b$ entonces $-b \leqq -x \leqq -a$ por lo que $|x| \leqq \alpha$.*

Como $n > N_0$,entonces $n > 4\alpha^2\frac{p}{q}$, así pues $\alpha\sqrt{\frac{p}{nq}} < \frac{1}{2}$. De la definición de valor absoluto se sabe que

$$x \leqq |x| \text{ y } -x \leqq |x|.$$

Entonces, de la observación 11 y de la definición de valor absoluto podemos establecer que

$$-x\sqrt{\frac{p}{nq}} \leqq |x|\sqrt{\frac{p}{nq}} \leqq \alpha\sqrt{\frac{p}{nq}} < \frac{1}{2}.$$

de donde

$$-\frac{1}{2} < x\sqrt{\frac{p}{nq}},$$

sumando 1 de ambos lados de la desigualdad

$$1 - \frac{1}{2} < 1 + x\sqrt{\frac{p}{nq}},$$

por lo tanto

$$1 + x\sqrt{\frac{q}{np}} > 0.$$

Por otro lado $n > N_0$,entonces $n > 4\alpha^2\frac{q}{p}$, así pues $\alpha\sqrt{\frac{q}{np}} < \frac{1}{2}$. Análogamente de la definición de valor absoluto y de la observación 11 se tiene que

$$x\sqrt{\frac{q}{np}} \leqq |x|\sqrt{\frac{q}{np}} \leqq \alpha\sqrt{\frac{q}{np}} < \frac{1}{2},$$

de donde

$$1 - x\sqrt{\frac{p}{nq}} > \frac{1}{2},$$

por lo que

$$1 - x\sqrt{\frac{p}{nq}} > 0.$$

Así pues, el cuarto paso ha quedado demostrado.

Paso 5 *Si* $n > N_0$, $0 \leqq m \leqq n$ *y* $x \in [a, b]$ *entonces*

$$\log\left(1 + x\sqrt{\frac{q}{np}}\right) = x\sqrt{\frac{q}{np}} - \frac{\left(x\sqrt{\frac{q}{np}}\right)^2}{2} + \frac{1}{3!}\left(x\sqrt{\frac{q}{np}}\right)^3 \frac{2}{(1+\gamma)^3}$$

en donde $\frac{8}{27} < \frac{1}{(1+\gamma)^3} < 8$. *Y*

$$\log\left(1 - x\sqrt{\frac{p}{nq}}\right) = -x\sqrt{\frac{p}{nq}} - \frac{1}{2}\frac{p}{nq}x^2 + \frac{1}{3}x^3\left(\frac{p}{nq}\right)^{\frac{3}{2}} \frac{1}{(1+\varphi)^3}$$

en donde $\frac{8}{27} < \frac{1}{(1+\varphi)^3} < 8$.

Demostración *El teorema de Taylor establece que si* $f : [a, b] \to R$ *de clase* $C^{n+1}[a, b]$. *Existe* $\theta \in (a, b)$ *tal que*

$$f(b) = \sum_{i=0}^{n} \frac{1}{i!} f^{(i)}(a)(b-a)^i + f^{(n+1)}(\theta)\frac{(b-a)^{n+1}}{(n+1)!}.$$

Aplicando el teorema mencionado a la función $f(z) = \log(1+z)$ *con* $|z| < 1$, *entonces existe* γ *entre* 0 *y* z, *tal que*

$$f(z) = 0 + \frac{1}{1!}(1)z + \frac{1}{2!}(-1)z^2 + \frac{2}{3!}\frac{1}{(1+\gamma)^3}z^3.$$

Ahora se consideran las hipótesis, puesto que $x \in [a, b]$ *y* $n > 4\alpha^2\frac{p}{q}$ *entonces por el paso 4* $x\sqrt{\frac{q}{np}} < \frac{1}{2}$,*en consecuencia*

$$\log\left(1 + x\sqrt{\frac{q}{np}}\right) = x\sqrt{\frac{q}{np}} - \frac{1}{2}\frac{q}{np}x^2 + \frac{1}{3}x^3\left(\frac{q}{np}\right)^{\frac{3}{2}} \frac{1}{(1+\gamma)^3},$$

con γ *entre* 0 *y* $x\sqrt{\frac{q}{np}}$.

Para acotar $\frac{1}{(1+\gamma)^3}$ *se observa por el paso 4 que* $|x|\sqrt{\frac{q}{np}} < \frac{1}{2}$, *y además* γ *entre* 0 *y* $x\sqrt{\frac{q}{np}}$, *de donde*

$$\begin{aligned} -\frac{1}{2} &< \gamma < \frac{1}{2} \\ \frac{1}{2} &< \gamma + 1 < \frac{3}{2} \\ 2 &< \frac{1}{\gamma+1} < \frac{2}{3} \\ \frac{8}{27} &< \frac{1}{(1+\gamma)^3} < 8. \end{aligned}$$

Análogamente si $x \in [a,b]$ *y* $n > 4\alpha^2 \frac{q}{p}$ *entonces por el paso 4* $x\sqrt{\frac{p}{nq}} < \frac{1}{2}$,*entonces*

$$\log\left(1 - x\sqrt{\frac{p}{nq}}\right) = -x\sqrt{\frac{p}{nq}} - \frac{1}{2}\frac{p}{nq}x^2 + \frac{1}{3}x^3\left(\frac{p}{nq}\right)^{\frac{3}{2}}\frac{1}{(1+\varphi)^3},$$

con $0 < \varphi < x\sqrt{\frac{p}{nq}}$.*y análogamente para acotar* φ, *se tiene que* $|x|\sqrt{\frac{p}{nq}} < \frac{1}{2}$ *y* φ *está entre* 0 *y* $x\sqrt{\frac{p}{nq}}$ *entonces*

$$\frac{8}{27} < \frac{1}{(1+\varphi)^3} < 8.$$

Por lo que el paso ha quedado demostrado.

Paso 6 *Sea* $a \leqq x \leqq b$, *entonces* $np\left(1 + a\sqrt{\frac{q}{np}}\right) \leqq m$ *y* $nq\left(1 - a\sqrt{\frac{p}{nq}}\right) \leqq n - m$.

Demostración *Por hipótesis* $a \leqq x \leqq b$, *entonces se tiene que*

$$a \leqq \frac{m - np}{\sqrt{npq}} \leqq b$$

en consecuencia,

$$a\sqrt{npq} + np \leqq m \leqq b\sqrt{npq} + np$$

y factorizando np *se tiene,*

$$np\left(1 + a\sqrt{\frac{q}{np}}\right) \leqq m \leqq np\left(1 + b\sqrt{\frac{q}{np}}\right),$$

por lo pronto, sólo se utiliza la desigualdad de la izquierda

$$np\left(1 + a\sqrt{\frac{q}{np}}\right) \leqq m. \tag{3.9}$$

Análogamente si $-a \leqq -x \leqq -b$, *entonces*

$$nq\left(1 - a\sqrt{\frac{p}{nq}}\right) \leqq n - m \leqq nq\left(1 - b\sqrt{\frac{p}{nq}}\right),$$

aquí también, sólo se utiliza la desigualdad de la izquierda

$$nq\left(1-a\sqrt{\frac{p}{nq}}\right)\leqq n-m. \tag{3.10}$$

Por lo que el paso 6 ha quedado demostrado.

■

Ahora se pueden acotar los términos R, S y θ.

Teorema 12 ***Demostración.***

Paso 7 *Sea $\eta>0$ y sea N_1 un natural tal que*

$$N_1>\frac{1}{\eta^2}\left(\frac{3\alpha}{2}\sqrt{\frac{q}{p}}+\frac{3\alpha}{2}\sqrt{\frac{p}{q}}+\frac{3q}{4p}\alpha^2+\frac{3p}{4q}\alpha^2+4\alpha^3\left(\frac{q}{p}\right)^{\frac{3}{2}}+4\alpha^3\left(\frac{p}{q}\right)^{\frac{3}{2}}\right)^2,$$

si $n>\text{máx}(N_0,N_1)$, $0\leqq m\leqq n$ y $x\in[a,b]$, entonces $|S|<\frac{\eta}{3}$.

Demostración *Como $n>N_1$ implica lo siguiente:*

$$\begin{aligned}
n &> \frac{1}{\eta^2}\left(\frac{3\alpha}{2}\sqrt{\frac{q}{p}}+\frac{3\alpha}{2}\sqrt{\frac{p}{q}}+\frac{3q}{4p}\alpha^2+\frac{3p}{4q}\alpha^2+4\alpha^3\left(\frac{q}{p}\right)^{\frac{3}{2}}+4\alpha^3\left(\frac{p}{q}\right)^{\frac{3}{2}}\right)^2\\
\sqrt{n} &> \frac{1}{\eta}\left(\frac{3\alpha}{2}\sqrt{\frac{q}{p}}+\frac{3\alpha}{2}\sqrt{\frac{p}{q}}+\frac{3q}{4p}\alpha^2+\frac{3p}{4q}\alpha^2+4\alpha^3\left(\frac{q}{p}\right)^{\frac{3}{2}}+4\alpha^3\left(\frac{p}{q}\right)^{\frac{3}{2}}\right)=\\
&= \frac{3}{\eta}\left(\frac{\alpha}{2}\sqrt{\frac{q}{p}}+\frac{\alpha}{2}\sqrt{\frac{p}{q}}+\frac{q}{4p}\alpha^2+\frac{p}{4q}\alpha^2+\frac{4}{3}\alpha^3\left(\frac{q}{p}\right)^{\frac{3}{2}}+\frac{4}{3}\alpha^3\left(\frac{p}{q}\right)^{\frac{3}{2}}\right)
\end{aligned}$$

consecuentemente por ser números positivos se tiene que

$$\begin{aligned}
\frac{\eta}{3} &> \frac{1}{\sqrt{n}}\left(\frac{\alpha}{2}\sqrt{\frac{q}{p}}+\frac{\alpha}{2}\sqrt{\frac{p}{q}}+\frac{q}{4p}\alpha^2+\frac{p}{4q}\alpha^2+\frac{4}{3}\alpha^3\left(\frac{q}{p}\right)^{\frac{3}{2}}+\frac{4}{3}\alpha^3\left(\frac{p}{q}\right)^{\frac{3}{2}}\right)=\\
&= \frac{1}{\sqrt{n}}\left(\frac{\alpha}{2}\sqrt{\frac{q}{p}}+\frac{\alpha}{2}\sqrt{\frac{p}{q}}\right)+\frac{1}{\sqrt{n}}\left(\frac{q}{4p}\alpha^2+\frac{p}{4q}\alpha^2\right)+ \qquad (3.11)\\
&\quad +\frac{1}{\sqrt{n}}\left(\frac{4}{3}\alpha^3\left(\frac{q}{p}\right)^{\frac{3}{2}}+\frac{4}{3}\alpha^3\left(\frac{p}{q}\right)^{\frac{3}{2}}\right) \qquad (3.12)
\end{aligned}$$

se puede observar que si $n>1$ entonces se cumple que $\sqrt{n}<n<n\sqrt{n}$, de donde

$$\frac{1}{n\sqrt{n}}<\frac{1}{n}<\frac{1}{\sqrt{n}} \tag{3.13}$$

entonces de 3.11 y 3.13, se deduce que

$$\begin{aligned}
\frac{\eta}{3} &> \frac{1}{\sqrt{n}}\left(\frac{\alpha}{2}\sqrt{\frac{q}{p}}+\frac{\alpha}{2}\sqrt{\frac{p}{q}}\right)+\frac{1}{n}\left(\frac{q}{4p}\alpha^2+\frac{p}{4q}\alpha^2\right)+\frac{1}{n^{\frac{3}{2}}}\left(\frac{4}{3}\alpha^3\left(\frac{q}{p}\right)^{\frac{3}{2}}+\frac{4}{3}\alpha^3\left(\frac{p}{q}\right)^{\frac{3}{2}}\right)= \\
&= \frac{\alpha}{2}\sqrt{\frac{q}{np}}+\frac{1}{4}\frac{q}{np}\alpha^2+\frac{8}{6}\alpha^3\left(\frac{q}{np}\right)^{\frac{3}{2}}+\frac{\alpha}{2}\sqrt{\frac{p}{nq}}+\frac{1}{4}\frac{p}{nq}\alpha^2+\frac{8}{6}\alpha^3\left(\frac{p}{nq}\right)^{\frac{3}{2}}.
\end{aligned}$$

Ahora, por la desigualdad del triángulo y por la observación 11, se puede establecer que

$$\begin{aligned}
\frac{\eta}{3} &> \left|\frac{x}{2}\sqrt{\frac{q}{np}}+\frac{1}{4}\frac{q}{np}x^2+\frac{8}{6}x^3\left(\frac{q}{np}\right)^{\frac{3}{2}}+\frac{x}{2}\sqrt{\frac{p}{nq}}+\frac{1}{4}\frac{p}{nq}x^2+\frac{8}{6}x^3\left(\frac{p}{nq}\right)^{\frac{3}{2}}\right| > \\
&> \left|-\frac{x}{2}\sqrt{\frac{q}{np}}+\frac{1}{4}\frac{q}{np}x^2-\frac{1}{6}x^3\left(\frac{q}{np}\right)^{\frac{3}{2}}\frac{1}{(1+\gamma)^3}+\frac{x}{2}\sqrt{\frac{p}{nq}}+\frac{1}{4}\frac{p}{nq}x^2-\frac{1}{6}x^3\left(\frac{p}{nq}\right)^{\frac{3}{2}}\frac{1}{(1+\varphi)^3}\right| \\
&= \left|-\frac{1}{2}\left[\begin{array}{c} x\sqrt{\frac{q}{np}}-\frac{1}{2}\frac{q}{np}x^2+\frac{1}{3}x^3\left(\frac{q}{np}\right)^{\frac{3}{2}}\frac{1}{(1+\gamma)^3}-x\sqrt{\frac{p}{nq}}- \\ -\frac{1}{2}\frac{p}{nq}x^2+\frac{1}{3}x^3\left(\frac{p}{nq}\right)^{\frac{3}{2}}\frac{1}{(1+\varphi)^3}\end{array}\right]\right|
\end{aligned}$$

y como se puede ver, la expresión anterior contiene los desarrollos de Taylor de $\log\left(1+x\sqrt{\frac{q}{np}}\right)$ *y* $\log\left(1-x\sqrt{\frac{p}{nq}}\right)$ *.Por lo tanto*

$$\frac{\eta}{3} > \left|-\frac{1}{2}\left[\log\left(1+x\sqrt{\frac{q}{np}}\right)+\log\left(1-x\sqrt{\frac{p}{nq}}\right)\right]\right|$$

y utilizando las propiedades de los logaritmos, se tiene que

$$\begin{aligned}
\frac{\eta}{3} &> \left|-\frac{1}{2}\log\left[\left(1+x\sqrt{\frac{q}{np}}\right)\left(1-x\sqrt{\frac{p}{nq}}\right)\right]\right| \\
&= \left|\log 1-\log\left[\left(1+x\sqrt{\frac{q}{np}}\right)\left(1-x\sqrt{\frac{p}{nq}}\right)\right]^{\frac{1}{2}}\right| \\
&= \left|\log\frac{1}{\sqrt{\left(1+x\sqrt{\frac{q}{np}}\right)\left(1-x\sqrt{\frac{p}{nq}}\right)}}\right|, \qquad (3.14)
\end{aligned}$$

aprovechando que se puede escribir a $\sqrt{\frac{n}{m(n-m)}}$ *de la siguiente forma:*

$$\begin{aligned}
\sqrt{\frac{n}{m(n-m)}} &= \sqrt{\frac{n}{n^2pq}}\sqrt{\left(\frac{np}{m}\right)\left(\frac{nq}{n-m}\right)} \\
&= \sqrt{\frac{1}{npq}}\frac{1}{\sqrt{\left(\frac{m}{np}\right)\left(\frac{n-m}{nq}\right)}} \\
&= \frac{1}{\sqrt{npq}}\frac{1}{\sqrt{\left(1+x\sqrt{\frac{q}{np}}\right)\left(1-x\sqrt{\frac{p}{nq}}\right)}},
\end{aligned}$$

entonces 3.14, se puede expresar como:

$$\begin{aligned}
\frac{\eta}{3} &> \left|\log\frac{1}{\sqrt{\left(1+x\sqrt{\frac{q}{np}}\right)\left(1-x\sqrt{\frac{p}{nq}}\right)}}\right| = \\
&= \left|\log\left(\sqrt{\frac{n}{m(n-m)}}\sqrt{npq}\right)\right| \\
&= \left|\log\sqrt{\frac{n}{m(n-m)}}+\log\sqrt{npq}\right| = |S|
\end{aligned}$$

Por lo tanto $|S| < \frac{\eta}{3}$.

Paso 8 *Sea* $\eta > 0$ *y sea* N_2 *un natural tal que*

$$N_2 > \frac{1}{\eta^2}\left(\frac{169}{4}\alpha^6\frac{p^3}{q}+\frac{169}{4}\alpha^6\frac{q^3}{p}+8\alpha^8\frac{p^4}{q}+8\alpha^8\frac{q^4}{p}\right),$$

si $n > \text{máx}(N_0, N_2)$, $0 \leqq m \leqq n$ *y* $x \in [a,b]$, *entonces* $|R| < \frac{\eta}{3}$.

Demostración *Como* $n > N_2$ *implica lo siguiente*

$$\begin{aligned}
n &> \frac{1}{\eta^2}\left(\frac{169}{4}\alpha^6\frac{p^3}{q}+\frac{169}{4}\alpha^6\frac{q^3}{p}+8\alpha^8\frac{p^4}{q}+8\alpha^8\frac{q^4}{p}\right) \\
\sqrt{n} &> \frac{1}{\eta}\left(\frac{13}{2}\alpha^3\frac{p^{\frac{3}{2}}}{\sqrt{q}}+\frac{13}{2}\alpha^3\frac{q^{\frac{3}{2}}}{\sqrt{p}}+8\alpha^4\frac{p^2}{\sqrt{q}}+8\alpha^4\frac{q^2}{\sqrt{p}}\right) \\
\eta &> \frac{1}{\sqrt{n}}\left(\frac{13}{2}\alpha^3\frac{p^{\frac{3}{2}}}{\sqrt{q}}+\frac{13}{2}\alpha^3\frac{q^{\frac{3}{2}}}{\sqrt{p}}+8\alpha^4\frac{p^2}{\sqrt{q}}+8\alpha^4\frac{q^2}{\sqrt{p}}\right) \\
\frac{\eta}{3} &> \frac{1}{\sqrt{n}}\left(\frac{13}{6}\alpha^3\frac{p^{\frac{3}{2}}}{\sqrt{q}}+\frac{13}{6}\alpha^3\frac{q^{\frac{3}{2}}}{\sqrt{p}}+\frac{8}{3}\alpha^4\frac{p^2}{\sqrt{q}}+\frac{8}{3}\alpha^4\frac{q^2}{\sqrt{p}}\right)
\end{aligned}$$

resolviendo el producto

$$\frac{\eta}{3} > \left(\frac{13}{6}\alpha^3\frac{p^{\frac{3}{2}}}{\sqrt{nq}} + \frac{13}{6}\alpha^3\frac{q^{\frac{3}{2}}}{\sqrt{np}} + \frac{8}{3}\alpha^4\frac{p^2}{\sqrt{nq}} + \frac{8}{3}\alpha^4\frac{q^2}{\sqrt{np}}\right).$$

Ahora, por la desigualdad del triángulo y por la observación 11, se puede establecer que

$$\begin{aligned}\frac{\eta}{3} &> \left|\frac{13}{6}x^3\frac{p^{\frac{3}{2}}}{\sqrt{nq}} - \frac{13}{6}x^3\frac{q^{\frac{3}{2}}}{\sqrt{np}} - \frac{8}{3}x^4\frac{p^2}{\sqrt{nq}} - \frac{8}{3}x^4\frac{q^2}{\sqrt{np}}\right| = \\ &= \left|\frac{8}{3}x^3\frac{p^{\frac{3}{2}}}{\sqrt{nq}} - \frac{1}{2}x^3\frac{p^{\frac{3}{2}}}{\sqrt{nq}} - \frac{8}{3}x^3\frac{q^{\frac{3}{2}}}{\sqrt{np}} + \frac{1}{2}x^3\frac{q^{\frac{3}{2}}}{\sqrt{np}} - \frac{8}{3}x^4\frac{p^2}{\sqrt{nq}} - \frac{8}{3}x^4\frac{q^2}{\sqrt{np}}\right|. \quad (3.15)\end{aligned}$$

Por el paso 5 que $\frac{8}{27} < \frac{1}{(1+\gamma)^3} < 8$ *y* $\frac{8}{27} < \frac{1}{(1+\varphi)^3} < 8$, *así que se puede ver a 3.15 como*

$$\begin{aligned}\frac{\eta}{3} &> \left|\frac{x^3p^{\frac{3}{2}}}{3\sqrt{nq}}\frac{1}{(1+\varphi)^3} - \frac{x^3p^{\frac{3}{2}}}{2\sqrt{nq}} - \frac{x^3q^{\frac{3}{2}}}{3\sqrt{np}}\frac{1}{(1+\gamma)^3} + \frac{x^3q^{\frac{3}{2}}}{2\sqrt{np}} - \frac{x^4p^2}{3\sqrt{nq}}\frac{1}{(1+\varphi)^3} - \frac{x^4q^2}{3\sqrt{np}}\frac{1}{(1+\gamma)^3}\right| = \\ &= \left|\frac{1}{2}x^3\left(\frac{q^{\frac{3}{2}}}{\sqrt{np}} - \frac{p^{\frac{3}{2}}}{\sqrt{nq}}\right) + \frac{1}{3(1+\varphi)^3}\left(\frac{x^3p^{\frac{3}{2}}}{\sqrt{nq}} - \frac{x^4p^2}{\sqrt{nq}}\right) - \frac{1}{3(1+\gamma)^3}\left(\frac{x^3q^{\frac{3}{2}}}{\sqrt{np}} + \frac{x^4q^2}{\sqrt{np}}\right)\right|,\end{aligned}$$

sumando un cero, se tiene

$$\frac{\eta}{3} > \left|\begin{array}{c}\frac{1}{2}x^3\left(\frac{q^{\frac{3}{2}}}{\sqrt{np}} - \frac{p^{\frac{3}{2}}}{\sqrt{nq}}\right) + \frac{1}{3(1+\varphi)^3}\left(\frac{x^3p^{\frac{3}{2}}}{\sqrt{nq}} - \frac{x^4p^2}{\sqrt{nq}}\right) - \frac{1}{3(1+\gamma)^3}\left(\frac{x^3q^{\frac{3}{2}}}{\sqrt{np}} + \frac{x^4q^2}{\sqrt{np}}\right) + \\ +\frac{x^2}{2}(p+q) - x^2(p+q) + \frac{x^2}{2} + x\sqrt{npq} - x\sqrt{npq}\end{array}\right|$$

reacomodando términos queda la expresión

$$\frac{\eta}{3} > \left|\begin{array}{c}-x\sqrt{npq} + \frac{1}{2}x^2q - \frac{1}{3}x^3\left(\frac{q^{\frac{3}{2}}}{\sqrt{np}}\right)\frac{1}{(1+\gamma)^3} - x^2q + \frac{1}{2}x^3\left(\frac{q^{\frac{3}{2}}}{\sqrt{np}}\right) - \frac{1}{3}x^4\left(\frac{q^2}{\sqrt{np}}\right)\frac{1}{(1+\gamma)^3} + \\ +x\sqrt{npq} + \frac{1}{2}x^2p + \frac{1}{3}x^3\left(\frac{p^{\frac{3}{2}}}{\sqrt{nq}}\right)\frac{1}{(1+\varphi)^3} - x^2p - \frac{1}{2}x^3\left(\frac{p^{\frac{3}{2}}}{\sqrt{nq}}\right) - \frac{1}{3}x^4\left(\frac{p^2}{\sqrt{nq}}\right)\frac{1}{(1+\varphi)^3} + \frac{x^2}{2}\end{array}\right|$$

factorizando respectivamente $\left(-np - x\sqrt{npq}\right)$ *y* $-\left(np - x\sqrt{npq}\right)$, *se tiene*

$$\frac{\eta}{3} > \left|\begin{array}{c}\left(-np - x\sqrt{npq}\right)\left[x\sqrt{\frac{q}{np}} - \frac{1}{2}\frac{x^2q}{np} + \frac{1}{3}x^3\left(\frac{q}{np}\right)^{\frac{3}{2}}\frac{1}{(1+\gamma)^3}\right] - \\ -\left(np - x\sqrt{npq}\right)\left[x\sqrt{\frac{q}{np}} - \frac{1}{2}\frac{x^2q}{np} + \frac{1}{3}x^3\left(\frac{q}{np}\right)^{\frac{3}{2}}\frac{1}{(1+\gamma)^3}\right] + \frac{x^2}{2}\end{array}\right| \quad (3.16)$$

por el paso 5, se sabe que las expresiones dentro del corchete corresponden a los desarrollos de Taylor de $\log\left(1 + x\sqrt{\frac{q}{np}}\right)$ *y* $\log\left(1 - x\sqrt{\frac{p}{nq}}\right)$, *así que 3.16 se puede expresar de la siguiente forma*

$$\begin{aligned}\frac{\eta}{3} &> \left|(-np - x\sqrt{npq})\log\left(1 + x\sqrt{\frac{q}{np}}\right) - (np - x\sqrt{npq})\log\left(1 - x\sqrt{\frac{p}{nq}}\right) + \frac{x^2}{2}\right| = \\ &= \left|-(np + x\sqrt{npq})\log\left(\frac{np + x\sqrt{npq}}{np}\right) - (nq - x\sqrt{npq})\log\left(\frac{nq - x\sqrt{npq}}{nq}\right) + \frac{x^2}{2}\right|.\end{aligned}$$

Por hipótesis $a \leqq x \leqq b$, por lo tanto

$$\frac{\eta}{3} > \left| -\left(np + a\sqrt{npq}\right)\log\left(\frac{np + a\sqrt{npq}}{np}\right) - \left(nq - a\sqrt{npq}\right)\log\left(\frac{nq - a\sqrt{npq}}{nq}\right) + \frac{x^2}{2} \right|$$

y por el paso 6 $-\left(np + a\sqrt{npq}\right) \geqq -m$ *y* $-\left(nq - a\sqrt{npq}\right) \geqq -(n-m)$, *de modo que*

$$\begin{aligned} \frac{\eta}{3} &> \left| -m\log\left(\frac{m}{np}\right) - (n-m)\log\left(\frac{n-m}{nq}\right) + \frac{x^2}{2} \right| = \\ &= \left| \log\left(\frac{m}{np}\right)^{-m} + \log\left(\frac{n-m}{nq}\right)^{-(n-m)} + \frac{x^2}{2} \right| \\ &= \left| \log\left(\frac{np}{m}\right)^{m} + \log\left(\frac{nq}{n-m}\right)^{n-m} + \frac{x^2}{2} \right| \\ &= \left| \log\left[\left(\frac{np}{m}\right)^{m}\left(\frac{nq}{n-m}\right)^{n-m}\right] + \frac{x^2}{2} \right| = |R| \end{aligned}$$

Por lo tanto $|R| < \frac{\eta}{3}$.

Paso 9 *Si $x \in [a,b]$,sea $\eta > 0$ y N_3 un natural tal que*

$$N_3 > \frac{1}{4\eta}\left(1 + \frac{2}{p} + \frac{2}{q}\right),$$

si $n > \text{máx}(N_0, N_3)$, $0 \leqq m \leqq n$ y $x \in [a,b]$, entonces $|\theta| < \frac{\eta}{3}$.

Demostración *Si $n > N_0$ entonces por el paso 4 $|x|\sqrt{\frac{q}{np}} < \frac{1}{2}$ y por hipótesis $a \leqq x \leqq b$, por lo que*

$$-\frac{1}{2} < a\sqrt{\frac{q}{np}} < \frac{1}{2}$$

de donde

$$\frac{1}{2} < 1 + a\sqrt{\frac{q}{np}} < \frac{3}{2}$$

y tomando sus recíprocos

$$\frac{3}{2} < \frac{1}{1 + a\sqrt{\frac{q}{np}}} < 2. \tag{3.17}$$

Análogamente si $|x|\sqrt{\frac{p}{nq}} < \frac{1}{2}$, entonces

$$\frac{3}{2} < \frac{1}{1 - a\sqrt{\frac{p}{nq}}} < 2. \tag{3.18}$$

Así pues de las desigualdades 3.9 y 3.10 del paso 6, se toman los recíprocos y se tienen las dos expresiones siguientes

$$\frac{1}{m} \leqq \frac{1}{np\left(1+a\sqrt{\frac{q}{np}}\right)} \tag{3.19}$$

y

$$\frac{1}{n-m} \leqq \frac{1}{nq\left(1-a\sqrt{\frac{p}{nq}}\right)}. \tag{3.20}$$

Sumando 3.19 y 3.20 se tiene que

$$\frac{1}{m}+\frac{1}{n-m} \leqq \frac{1}{np\left(1+a\sqrt{\frac{q}{np}}\right)}+\frac{1}{nq\left(1-a\sqrt{\frac{p}{nq}}\right)},$$

factorizando $\frac{1}{n}$

$$\frac{1}{m}+\frac{1}{n-m} \leqq \frac{1}{n}\left(\frac{1}{p\left(1+a\sqrt{\frac{q}{np}}\right)}+\frac{1}{q\left(1-a\sqrt{\frac{p}{nq}}\right)}\right),$$

ahora, sumando $\frac{1}{n}$ *de ambos lados de la desigualdad*

$$\frac{1}{n}+\frac{1}{m}+\frac{1}{n-m} \leqq \frac{1}{n}+\frac{1}{n}\left(\frac{1}{p\left(1+a\sqrt{\frac{q}{np}}\right)}+\frac{1}{q\left(1-a\sqrt{\frac{p}{nq}}\right)}\right)$$

multiplicando $\frac{1}{12}$ *de ambos lados y factorizando* $\frac{1}{12n}$ *en el lado derecho de la desigualdad, se tiene*

$$\frac{1}{12}\left(\frac{1}{n}+\frac{1}{m}+\frac{1}{n-m}\right) \leqq \frac{1}{12n}\left(1+\frac{1}{p\left(1+a\sqrt{\frac{q}{np}}\right)}+\frac{1}{q\left(1-a\sqrt{\frac{p}{nq}}\right)}\right). \tag{3.21}$$

Por otro lado, si $n > N_3$ *entonces*

$$n > \frac{1}{4\eta}\left(1+\frac{2}{p}+\frac{2}{q}\right) \tag{3.22}$$

de donde

$$\frac{\eta}{3} > \frac{1}{12n}\left(1+\frac{2}{p}+\frac{2}{q}\right) \tag{3.23}$$

entonces de las desigualdades 3.17, 3.18 y 3.21 se deduce que

$$\frac{1}{12}\left(\frac{1}{n}+\frac{1}{m}+\frac{1}{n-m}\right) < \frac{1}{12n}\left(1+\frac{2}{p}+\frac{2}{q}\right) < \frac{\eta}{3}.$$

Por la fórmula de Stirling

$$|\theta| \leqq \frac{1}{12}\left(\frac{1}{n}+\frac{1}{m}+\frac{1}{n-m}\right) < \frac{1}{12n}\left(1+\frac{2}{p}+\frac{2}{q}\right) < \frac{\eta}{3}$$

y por lo tanto $|\theta| < \frac{\eta}{3}$.

A continuación se concluye la demostración.

Paso 10 *Sea* $0 < p < 1$ *y* $q = 1-p$, *y sean* $a < b$ *dos reales. Entonces para todo* $\varepsilon > 0$ *existe un natural* N_ε *tal que si* $n > N_\varepsilon$ *se tiene que, cuando* $0 \leqq m \leqq n$ *son tales que* $x_{(n,m)} \in [a,b]$, *entonces*

$$\left|\frac{\sqrt{npq}P_n(m)}{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}-1\right| < \varepsilon.$$

Demostración *Sea* $\varepsilon > 0$, *construyamos* $\eta = \log(1+\varepsilon) > 0$ *y sea*

$$N_\varepsilon = \max\{N_0, N_1, N_2, N_3\},$$

en donde N_0 *es como en el paso 4,* N_1 *como en el paso 7,* N_2 *como en el paso 8,* N_3 *como en el paso 9. Probaremos que* N_ε *cumple la propiedad deseada. Tomemos una* $n > N_\varepsilon$ *y* m *tal que* $0 \leqq m \leqq n$, *y además que* $x_{(n,m)} \in [a,b]$. *Entonces por el paso 1*

$$\left|\frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}}-1\right| < \varepsilon.$$

Ahora por el paso 3, se puede expresar a

$$\log\frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}} = S+R+\theta,$$

así pues, si $n > \max(N_0, N_1)$ *entonces* $|S| < \frac{\eta}{3}$; *si* $n > \max(N_0, N_2)$ *entonces* $|R| < \frac{\eta}{3}$ *y por último si* $n > N_3$, *entonces por el paso 9,* $|\theta| < \frac{\eta}{3}$.

$$\left|\log\frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}}{\sqrt{npq}}}\right| = |S+R+\theta| \leqq |S|+|R|+|\theta| < \frac{\eta}{3}+\frac{\eta}{3}+\frac{\eta}{3} = \eta,$$

como se deseaba.

■

Aplicaciones

En esta subsección, se desea analizar el significado del teorema, para ello se realizarán varios tipos de ejemplos, algunos serán calculos sobre casos particulares para estimar el tamaño de la n. Además se resolverá el ejemplo 18 expuesto en página 37 de la sección anterior, se resuelve haciendo uso del teorema y por otro lado se hacen los cálculos sin usar el teorema. Por último, se ofrece una interpretación geométrica para ver el tipo de aproximación a la que se puede llegar.

Ejemplo 1 Cada que ε varía, N varía. Así que dada la ε se puede decidir qué tan grande debe ser n para garantizar que

$$\left| \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\sqrt{npq}}} - 1 \right| < \varepsilon.$$

En la demostración vimos que para este caso, $N_\varepsilon = \text{máx}\{N_0, N_1, N_2, N_3\}$. Usando este hecho y con ayuda de Maple hemos realizado varios cálculos a partir de diferentes valores de ε para los cuales hemos obtenido el valor de N y podemos decir que para ε igual a 1, 0,1, 0,01 y para 0,001, N_ε puede tomar aproximadamente los valores $584, 656{,}5468$; $30, 922, 397{,}57$; $2, 837, 112, 753$; y $281, 180, 548, 300$, respectivamente. Como estamos interesados únicamente en los valores enteros de n, podemos deshechar la parte fraccionaria de N. En la siguiente tabla se concentran los valores que genera la expresión del teorema si $p = 0{,}5$ y $x_{(n,m)} \in [-4, 4]$.

ε	$n > N\varepsilon$		
$\varepsilon = 1$	$n > 584, 656$	implica	$\left| \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\sqrt{npq}}} - 1 \right| < 1$
$\varepsilon = 0{,}1$	$n > 30, 922, 397$	implica	$\left| \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\sqrt{npq}}} - 1 \right| < 0{,}1$
$\varepsilon = 0{,}01$	$n > 2, 837, 112, 753$	implica	$\left| \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\sqrt{npq}}} - 1 \right| < 0{,}01$
$\varepsilon = 0{,}001$	$n > 281, 180, 548, 300$	implica	$\left| \frac{P_n(m)}{\frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}{\sqrt{npq}}} - 1 \right| < 0{,}001$

Como se puede notar, los valores de n que se han obtenido son muy grandes, esto implicaría que las poblaciones también deben ser muy grandes para poder tener cierto grado de exactitud. Lo anterior se debe principalmente a que al hacer la demostración y acotar los términos no nos preocupamos por que las cotas tuvieran la característica de ser las más chicas, y como es bien sabido, los métodos para encontrar cotas que sean óptimas, requieren del uso de técnicas más sofisticadas y por el momento consideramos que ha sido un buen paso dar la ε y a partir de ella obtener la n.

Por otro lado, hay que reconocer que estos valores de n no son para nada didácticos, sin embargo hay otro tipo de cálculos que se pueden hacer para mostrar que las n no deben ser tan grandes, para que la expresión del teorema se cumpla.

En el ejemplo siguiente se muestran algunos cálculos para diferentes valores de n, así como diferentes valores de p y q.

Ejemplo 2 Resolveremos el ejemplo 18 expuesto en página 37, convenimos en que era necesario determinar el valor de $P_n(m)$ para $n = 10,000$ $m = 40$ y $p = 0{,}005$. Por el teorema de De Moivre-Laplace, que se demostró en la página 38, podemos decir que el cálculo de $P_n(m)$ se aproxima a la expresión

$$P_n(m) \sim \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{m-np}{\sqrt{npq}}\right)^2}}{\sqrt{npq}}.$$

Y sustituyendo los valores para el ejemplo

$$\sqrt{npq} = \sqrt{10000 \cdot 0{,}005 \cdot 0{,}995} = \sqrt{49{,}75} \approx 7{,}05,$$

de donde

$$\frac{m-np}{\sqrt{npq}} \approx -1{,}417 \approx -1{,}42$$

consecuentemente

$$P_n(m) \approx \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(-1{,}42)^2}}{7{,}05} \approx \frac{0{,}1456}{7{,}05} \approx 0{,}02064739431$$

Sin embargo, con ayuda de Maple podemos realizar los calculos para obtener el valor exacto y sin usar el teorema de DeMoivre-Laplace, de donde se tiene:

$$\begin{aligned} P_n(m) &= \binom{10000}{40} (0{,}005)^{40} (0{,}995)^{10000-40} \\ &\approx 0{,}02143481169 \end{aligned}$$

Como se puede ver la aproximación varía por algunas milésimas, esta variación tiene que ver también con el tipo de datos usados, ya que el valor de p es muy cercano a 0.

Se compara ahora usando los siguientes datos $P_n(m)$ para $n = 10,000$ $m = 5000$ y $p = 0{,}5$.

$$P_n(m) \approx \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(0)^2}}{50} \approx \frac{0{,}3989423}{50} \approx 0{,}007978845600$$

y sin usar el teorema

$$\begin{aligned} P_n(m) &= \binom{10000}{5000} (0{,}5)^{5000} (0{,}5)^{5000} \\ &\approx 0{,}007978646141. \end{aligned}$$

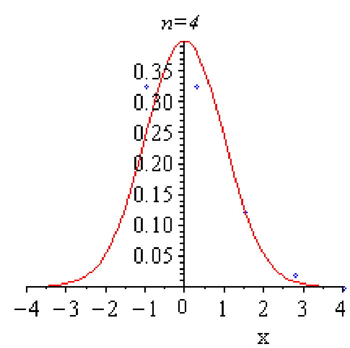
La aproximación es más exacta ya que p no es muy cercana ni a cero ni a uno.

Ejemplo 3

Para ilustrar el tipo de aproximaciones dadas por el teorema de DeMoivre-Laplace, y para ilustrar la interpretación geométrica de las transformaciones analíticas usadas en la prueba, se discute el siguiente ejemplo. Sea $p = 0{,}2$, y sea $n = 4$, en la siguiente tabla se proponen los valores de m, y se calcula x, $P_n(m)$ y $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

CASO n=4							
m	$\sqrt{npq}$	x	C_m^n	$P_n(m)$	$\sqrt{npq}\,P_n(m)$	$\frac{x^2}{2}$	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
0	0.8	-1.00	1	0.4096	**0.3277**	0.5000	**0.2420**
1	0.8	0.25	4	0.4096	**0.3277**	0.0313	**0.3867**
2	0.8	1.50	6	0.1536	**0.1229**	1.1250	**0.1295**
3	0.8	2.75	4	0.0256	**0.0205**	3.7813	**0.0091**
4	0.8	4.00	1	0.0016	**0.0013**	8.0000	**0.0001**

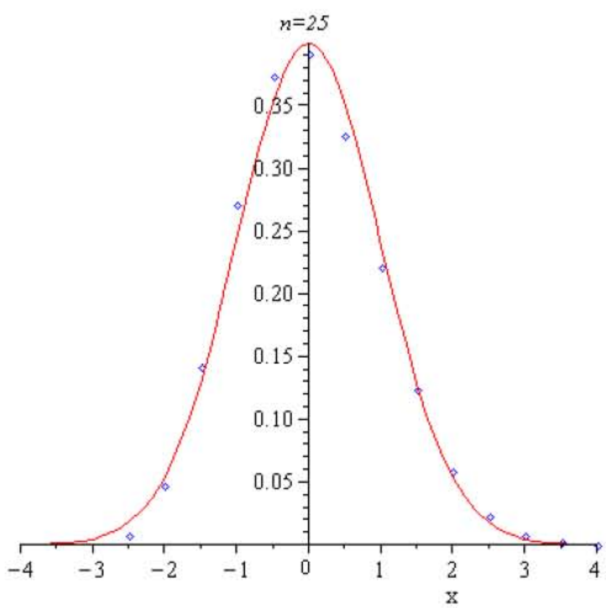
A continuación se grafican los puntos (x, y), considerando en el eje de las absisas los valores de x mostrados en la tabla, y en el eje de las ordenadas el valor de $y_n(m) = \sqrt{npq}P_n(m)$ junto con la función $y = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.



A continuación se lleva acabo el mismo procedimiento pero en diferentes casos $n = 25$, $n = 100$ y por último $n = 400$.

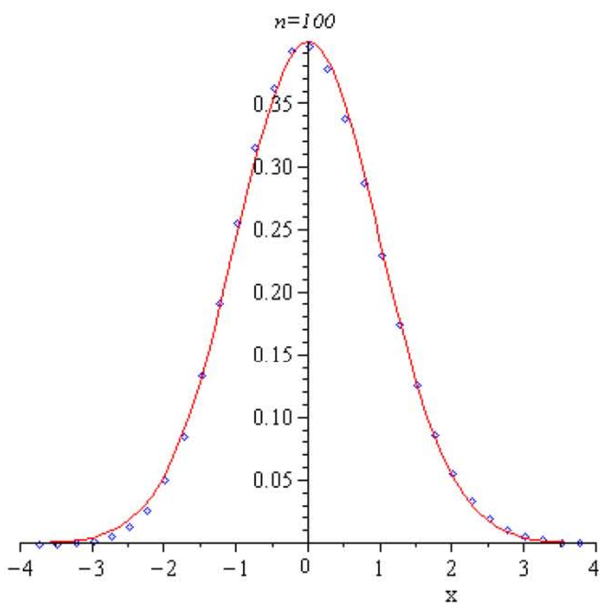
Caso n=25

CASO n=25							
m	$\sqrt{npq}$	x	C_m^n	$P_n(m)$	$\sqrt{npq}\,P_n(m)$	$\frac{x^2}{2}$	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
0	2	-2.50	1	0.0038	**0.0076**	3.1250	**0.0175**
1	2	-2.00	25	0.0236	**0.0472**	2.0000	**0.0540**
2	2	-1.50	300	0.0708	**0.1417**	1.1250	**0.1295**
3	2	-1.00	2300	0.1358	**0.2715**	0.5000	**0.2420**
4	2	-0.50	12650	0.1867	**0.3734**	0.1250	**0.3521**
5	2	0.00	53130	0.1960	**0.3920**	0.0000	**0.3989**
6	2	0.50	177100	0.1633	**0.3267**	0.1250	**0.3521**
7	2	1.00	480700	0.1108	**0.2217**	0.5000	**0.2420**
8	2	1.50	1081575	0.0623	**0.1247**	1.1250	**0.1295**
9	2	2.00	2042975	0.0294	**0.0589**	2.0000	**0.0540**
10	2	2.50	3268760	0.0118	**0.0236**	3.1250	**0.0175**
11	2	3.00	4457400	0.0040	**0.0080**	4.5000	**0.0044**
12	2	3.50	5200300	0.0012	**0.0023**	6.1250	**0.0009**
13	2	4.00	5200300	0.0003	**0.0006**	8.0000	**0.0001**
14	2	4.50	4457400	0.0001	**0.0001**	10.1250	**0.0000**
15	2	5.00	3268760	0.0000	**0.0000**	12.5000	**0.0000**
16	2	5.50	2042975	0.0000	**0.0000**	15.1250	**0.0000**
17	2	6.00	1081575	0.0000	**0.0000**	18.0000	**0.0000**
18	2	6.50	480700	0.0000	**0.0000**	21.1250	**0.0000**
19	2	7.00	177100	0.0000	**0.0000**	24.5000	**0.0000**
20	2	7.50	53130	0.0000	**0.0000**	28.1250	**0.0000**



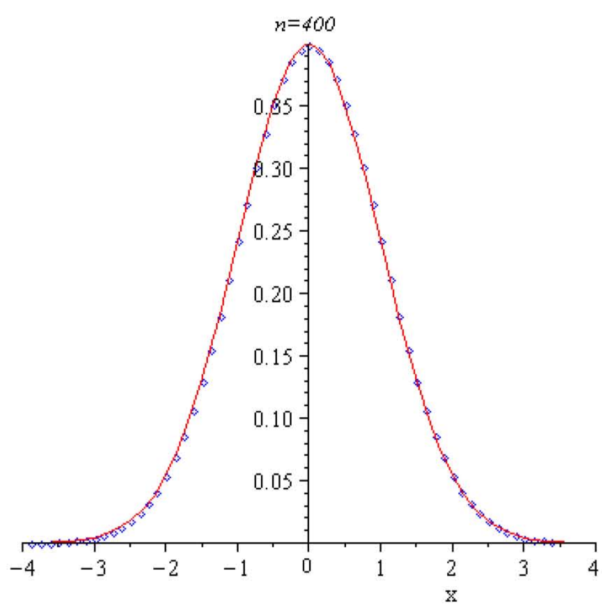
Caso n=100

CASO n=100							
m	$\sqrt{npq}$	x	C_m^n	$P_n(m)$	$\sqrt{npq}\,P_n(m)$	$\frac{x^2}{2}$	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
0	4	-5.00	1	0.0000	**0.0000**	12.5000	**0.0000**
1	4	-4.75	100	0.0000	**0.0000**	11.2813	**0.0000**
2	4	-4.50	4950	0.0000	**0.0000**	10.1250	**0.0000**
3	4	-4.25	161700	0.0000	**0.0000**	9.0313	**0.0000**
4	4	-4.00	3921225	0.0000	**0.0000**	8.0000	**0.0001**
5	4	-3.75	75287520	0.0000	**0.0001**	7.0313	**0.0004**
6	4	-3.50	1192052400	0.0001	**0.0002**	6.1250	**0.0009**
7	4	-3.25	16007560800	0.0002	**0.0008**	5.2813	**0.0020**
8	4	-3.00	1.86088E+11	0.0006	**0.0023**	4.5000	**0.0044**
9	4	-2.75	1.90223E+12	0.0015	**0.0059**	3.7813	**0.0091**
10	4	-2.50	1.73103E+13	0.0034	**0.0135**	3.1250	**0.0175**
11	4	-2.25	1.4163E+14	0.0069	**0.0275**	2.5313	**0.0317**
12	4	-2.00	1.05042E+15	0.0128	**0.0510**	2.0000	**0.0540**
13	4	-1.75	7.11054E+15	0.0216	**0.0863**	1.5313	**0.0863**
14	4	-1.50	4.41869E+16	0.0335	**0.1341**	1.1250	**0.1295**
15	4	-1.25	2.53338E+17	0.0481	**0.1922**	0.7813	**0.1826**
16	4	-1.00	1.34586E+18	0.0638	**0.2553**	0.5000	**0.2420**
17	4	-0.75	6.65013E+18	0.0789	**0.3154**	0.2813	**0.3011**
18	4	0.50	3.06645E+19	0.0909	**0.3636**	0.1250	**0.3521**
19	4	-0.25	1.32342E+20	0.0981	**0.3923**	0.0313	**0.3867**
20	4	0.00	5.35983E+20	0.0993	**0.3972**	0.0000	**0.3989**
21	4	0.25	2.04184E+21	0.0946	**0.3783**	0.0313	**0.3867**
22	4	0.50	7.33207E+21	0.0849	**0.3396**	0.1250	**0.3521**
23	4	0.75	2.48653E+22	0.0720	**0.2879**	0.2813	**0.3011**
24	4	1.00	7.97761E+22	0.0577	**0.2309**	0.5000	**0.2420**
25	4	1.25	2.42519E+23	0.0439	**0.1755**	0.7813	**0.1826**
26	4	1.50	6.99575E+23	0.0316	**0.1266**	1.1250	**0.1295**
27	4	1.75	1.91735E+24	0.0217	**0.0867**	1.5313	**0.0863**
28	4	2.00	4.99881E+24	0.0141	**0.0565**	2.0000	**0.0540**
29	4	2.25	1.24108E+25	0.0088	**0.0351**	2.5313	**0.0317**
30	4	2.50	2.93723E+25	0.0052	**0.0208**	3.1250	**0.0175**
31	4	2.75	6.63246E+25	0.0029	**0.0117**	3.7813	**0.0091**
32	4	3.00	1.43013E+26	0.0016	**0.0063**	4.5000	**0.0044**
33	4	3.25	2.94692E+26	0.0008	**0.0033**	5.2813	**0.0020**
34	4	3.50	5.80717E+26	0.0004	**0.0016**	6.1250	**0.0009**
35	4	3.75	1.09507E+27	0.0002	**0.0008**	7.0313	**0.0004**



Caso n=400

CASO n=400							
m	$\sqrt{npq}$	x	C_m^n	$P_n(m)$	$\sqrt{npq}\,P_n(m)$	$\frac{x^2}{2}$	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$
48	8	-4.00	3.37817E+62	0.0000	0.0001	8.0000	0.0001
49	8	-3.88	2.42677E+63	0.0000	0.0001	7.5078	0.0002
50	8	-3.75	1.70359E+64	0.0000	0.0002	7.0313	0.0004
51	8	-3.63	1.16913E+65	0.0000	0.0003	6.5703	0.0006
52	8	-3.50	7.84666E+65	0.0001	0.0005	6.1250	0.0009
53	8	-3.38	5.15215E+66	0.0001	0.0009	5.6953	0.0013
54	8	-3.25	3.31073E+67	0.0002	0.0014	5.2813	0.0020
55	8	-3.13	2.08275E+68	0.0003	0.0022	4.8828	0.0030
56	8	-3.00	1.28312E+69	0.0004	0.0034	4.5000	0.0044
57	8	-2.88	7.74377E+69	0.0006	0.0051	4.1328	0.0064
58	8	-2.75	4.5795E+70	0.0009	0.0076	3.7813	0.0091
59	8	-2.63	2.65456E+71	0.0014	0.0110	3.4453	0.0127
60	8	-2.50	1.50867E+72	0.0020	0.0156	3.1250	0.0175
61	8	-2.38	8.40901E+72	0.0027	0.0218	2.8203	0.0238
62	8	-2.25	4.59783E+73	0.0037	0.0298	2.5313	0.0317
63	8	-2.13	2.46677E+74	0.0050	0.0399	2.2578	0.0417
64	8	-2.00	1.29891E+75	0.0066	0.0526	2.0000	0.0540
65	8	-1.88	6.71436E+75	0.0085	0.0680	1.7578	0.0688
66	8	-1.75	3.40805E+76	0.0108	0.0862	1.5313	0.0863
67	8	-1.63	1.69894E+77	0.0134	0.1075	1.3203	0.1065
68	8	-1.50	8.31979E+77	0.0164	0.1316	1.1250	0.1295
69	8	-1.38	4.00315E+78	0.0198	0.1583	0.9453	0.1550
70	8	-1.25	1.89292E+79	0.0234	0.1871	0.7813	0.1826
71	8	-1.13	8.79806E+79	0.0272	0.2174	0.6328	0.2119
72	8	-1.00	4.02023E+80	0.0310	0.2483	0.5000	0.2420
73	8	-0.88	1.80635E+81	0.0349	0.2789	0.3828	0.2721
74	8	-0.75	7.98211E+81	0.0385	0.3082	0.2813	0.3011
75	8	-0.63	3.46956E+82	0.0419	0.3349	0.1953	0.3282
76	8	-0.50	1.48369E+83	0.0447	0.3580	0.1250	0.3521
77	8	-0.38	6.24306E+83	0.0471	0.3766	0.0703	0.3719
78	8	-0.25	2.58527E+84	0.0487	0.3899	0.0313	0.3867
79	8	-0.13	1.05374E+85	0.0497	0.3973	0.0078	0.3958
80	8	0.00	4.22814E+85	0.0498	0.3985	0.0000	0.3989
81	8	0.13	1.67038E+86	0.0492	0.3936	0.0078	0.3958
82	8	0.25	6.49817E+86	0.0478	0.3828	0.0313	0.3867
83	8	0.38	2.48966E+87	0.0458	0.3666	0.0703	0.3719
84	8	0.50	9.39551E+87	0.0432	0.3459	0.1250	0.3521
85	8	0.63	3.49292E+88	0.0402	0.3215	0.1953	0.3282
86	8	0.75	1.27938E+89	0.0368	0.2944	0.2813	0.3011
87	8	0.88	4.61754E+89	0.0332	0.2656	0.3828	0.2721
88	8	1.00	1.64238E+90	0.0295	0.2362	0.5000	0.2420
89	8	1.13	5.75754E+90	0.0259	0.2070	0.6328	0.2119
90	8	1.25	1.98955E+91	0.0224	0.1788	0.7813	0.1826
91	8	1.38	6.77759E+91	0.0190	0.1523	0.9453	0.1550
92	8	1.50	2.27639E+92	0.0160	0.1279	1.1250	0.1295
93	8	1.63	7.539E+92	0.0132	0.1059	1.3203	0.1065
94	8	1.75	2.46221E+93	0.0108	0.0865	1.5313	0.0863
95	8	1.88	7.9309E+93	0.0087	0.0696	1.7578	0.0688
96	8	2.00	2.51971E+94	0.0069	0.0553	2.0000	0.0540
97	8	2.13	7.89683E+94	0.0054	0.0433	2.2578	0.0417
98	8	2.25	2.44157E+95	0.0042	0.0335	2.5313	0.0317
99	8	2.38	7.44802E+95	0.0032	0.0255	2.8203	0.0238
100	8	2.50	2.24185E+96	0.0024	0.0192	3.1250	0.0175
101	8	2.63	6.65897E+96	0.0018	0.0143	3.4453	0.0127
102	8	2.75	1.95199E+97	0.0013	0.0105	3.7813	0.0091
103	8	2.88	5.64752E+97	0.0009	0.0076	4.1328	0.0064
104	8	3.00	1.6128E+98	0.0007	0.0054	4.5000	0.0044
105	8	3.13	4.54656E+98	0.0005	0.0038	4.8828	0.0030
106	8	3.25	1.2653E+99	0.0003	0.0026	5.2813	0.0020
107	8	3.38	3.4767E+99	0.0002	0.0018	5.6953	0.0013



Con estas gráficas se puede notar que la n no necesariamente debe ser tan grande, para que los datos se acerquen a la curva normal.

Por supuesto, en los ejemplos de esta sección, así como en cualquier otro poblema involucrado con la determinación de la probabilidad $P_n(m)$ para valores arbitrarios finitos de m y n que involucren el eorema de DeMoivre-Laplace, es necesario estimar el error que se genera al hacer la sustitución.

Por mucho tiempo, el teorema de DeMoivre Laplace fue aplicado a la solución de problemas de este tipo sin obtener resultados satisfactorios debido al error. La intuición denotaba que si la n era grande y el valor de p no era cercano a 0 o a 1, el uso del teorema arrojaba resultados satisfactorios. Actualmente existen muy buenas aproximaciones al error, de tal forma que se puede estimar para cualquier valor de n y p.

Teorema integral de DeMoivre-Laplace

El teorema Integral de DeMoivre-Laplace es una reformulación del teorema 3 de DeMoivre-Laplace que se presentó en la página 38. Esta sección está dedicada a la demostración de dicho teorema, la cual presentaremos con riguroso detalle. Este teorema es importante ya que será útil para poder enunciar el teorema central del limite que abordaremos en el último capítulo, y como veremos más adelante representa un papel muy importante en la teoría de la probabilidad.

Con respecto a la demostración, se puede comentar que primero se enuncian algunos lemas y corolarios que posteriormente servirán para la demostración del teorema Integral de DeMoivre-Laplace. Antes de enunciar el teorema, se establece la siguiente notación.

Sea μ "el número de veces que ocurre el evento A en n experimentos independientes", y sea $p = P(A)$ la probabilidad de dicho evento, donde $0 < p < 1$ y $q = 1 - p$. Para $a \leqq b$ reales, se define:

$$P_n[a, b) = P\left\{a \leqq \frac{\mu - np}{\sqrt{npq}} < b\right\}.$$

Para n arbitraria, se define $\Pi_n(x) : R \to R$, mediante:

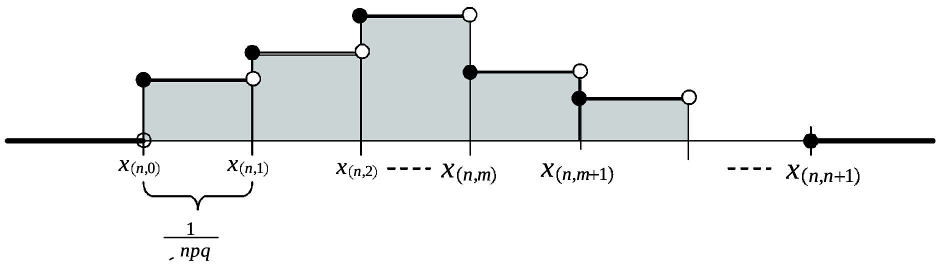
$$\Pi_n(x) = \begin{cases} 0 & \text{para } x < x_{(n,0)} = \frac{-np}{\sqrt{npq}} \\ \sqrt{npq}P_n(m) & \text{para } x_{(n,m)} \leqq x < x_{(n,m+1)} \text{ donde } m = 0, 1, ..., n \\ 0 & \text{para } x_{(n,n+1)} \leqq x \end{cases}$$

donde $x_{(n,n+1)} = x_{(n,n)} + \frac{1}{\sqrt{npq}} = \frac{nq+1}{\sqrt{npq}}$.

Y la probabilidad $P_n(m)$ es igual a $\sum P_n(m)$, tomada sobre esos valores de m para los cuales $a \leqq x_{(n,m)} < b$, donde como antes $x_{(n,m)} = \frac{m-np}{\sqrt{npq}}$. Es decir, igual al área limitada por la gráfica $\Pi_n(x)$, el eje x y las líneas $x = x_{(n,m)}$ y $x = x_{(n,m+1)}$, así se puede expresar $P_n(m)$ como sigue:

$$P_n(m) = \sqrt{npq}P_n(m)\left[x_{(n,m+1)} - x_{(n,m)}\right],$$

y geométricamente es la suma de todos los rectángulos.



Lema 13 *Existe N tal que $n > N$ entonces $\Pi_n(x) < \sqrt{\frac{2}{\pi}}$ para todo $x \in R$.*

Demostración. Se sabe que, si n es fijo, el valor máximo de $P_n(m)$ ocurre para $m_0 = [(n+1)\,p]$. Entonces, ya que

$$\begin{aligned}
\frac{1}{\sqrt{pq}} &\leqq -\frac{1}{\sqrt{npq}} < \frac{p-1}{\sqrt{npq}} = \frac{p-1+np-np}{\sqrt{npq}} = \\
&= \frac{(n+1)p-1-np}{\sqrt{npq}} < \frac{m_0-np}{\sqrt{npq}} < \frac{m_0+1-np}{\sqrt{npq}} \leqq \\
&\leqq \frac{(n+1)p+1-np}{\sqrt{npq}} = \frac{2}{\sqrt{npq}} \leqq \frac{2}{\sqrt{pq}},
\end{aligned}$$

se tiene que, puesto que $\Pi_n(x)$ se maximiza en el intervalo $\left[\frac{m_0-np}{\sqrt{npq}}, \frac{m_0+1-np}{\sqrt{npq}}\right)$, entonces también se maximiza en el intervalo $\left[-\frac{1}{\sqrt{pq}}, \frac{2}{\sqrt{pq}}\right)$, el cual no depende de n. Ahora, aplicamos al intervalo $\left[-\frac{1}{\sqrt{pq}}, \frac{2}{\sqrt{pq}}\right)$ el teorema Local de DeMoivre-Laplace y, si se hace $\varepsilon = 1$, se puede afirmar que existe N tal que si $n > N$, entonces si $-\frac{1}{\sqrt{pq}} \leqq x_{(n,m)} < \frac{2}{\sqrt{pq}}$, se tiene que

$$\left| \frac{\sqrt{npq}P_n(m)}{\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2_{(n,m)}}{2}}} - 1 \right| < 1,$$

entonces

$$\sqrt{npq}P_n(m) < \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2_{(n,m)}}{2}} \leqq \frac{2}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}}.$$

■

Lema 14 *Para cada $\varepsilon > 0$ existe N tal que si $n > N$ ocurre que, para a, b reales tales que $A \leqq a \leqq b \leqq B$ se tiene que*

$$\left| P_n\,[a,b) - \int_a^b \Pi_n(x)dx \right| < \varepsilon.$$

Demostración. Sea N_1 como en el Lema 13 y sea $\varepsilon > 0$. Existe N_2 tal que $n > N_2$ entonces $x_{(n,0)} < A$ y $B < x_{(n,n+1)}$. Si $n > \text{máx}(N_1, N_2)$ y sea $\overline{m}$ el máximo natural tal que $x_{(n,\underline{m})} < a$ y sea $\underline{m}$ el mínimo natural tal que $b \leqq x_{(n,\overline{m})}$. Entonces $0 \leqq \underline{m} < \overline{m} \leqq n+1$.Tenemos dos casos:

Caso 1 $\overline{m} = \underline{m}+1$

Entonces $P_n[a,b) = 0$ y por lo tanto

$$\begin{aligned}\left|P_n[a,b) - \int_a^b \Pi_n(x)dx\right| &= \left|\int_a^b \Pi_n(x)dx\right| \\ &\leqq \int_a^b |\Pi_n(x)|\,dx \leqq \sqrt{\frac{2}{\pi}}(b-a) < \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{npq}} < 2\sqrt{\frac{2}{\pi npq}}.\end{aligned}$$

Caso 2 $\underline{m}+1 < \overline{m}$

Entonces

$$\begin{aligned}P_n[a,b) &= \sum_{i=\underline{m}+1}^{\overline{m}-1} P_n(i) = \sum_{i=\underline{m}+1}^{\overline{m}-1} \sqrt{npq}P_n(i)\frac{1}{\sqrt{npq}} = \\ &= \int_{x_{(n,\underline{m}+1)}}^{x_{(n,\overline{m}-1)}} \Pi_n(x)dx = \left[\int_a^b - \int_a^{x_{(n,\underline{m}+1)}} - \int_{x_{(n,\overline{m}-1)}}^b\right]\Pi_n(x)dx\end{aligned}$$

de aquí se tiene que

$$\begin{aligned}\left|P_n[a,b) - \int_a^b \Pi_n(x)dx\right| &= \left|\int_a^{x_{(n,\underline{m}+1)}} \Pi_n(x)dx - \int_{x_{(n,\overline{m}-1)}}^b \Pi_n(x)dx\right| \\ &\leqq \int_a^{x_{(n,\underline{m}+1)}} |\Pi_n(x)|\,dx - \int_{x_{(n,\overline{m}-1)}}^b |\Pi_n(x)|\,dx \\ &\leqq \sqrt{\frac{2}{\pi}}\left(x_{(n,\underline{m}+1)} - a + b - x_{(n,\overline{m}-1)}\right) \\ &< \sqrt{\frac{2}{\pi}}\left(\frac{1}{\sqrt{npq}} + \frac{1}{\sqrt{npq}}\right) \\ &= 2\sqrt{\frac{2}{\pi npq}}.\end{aligned}$$

se toma entonces N_3 tal que $n > N_3$ entonces $2\sqrt{\frac{2}{\pi npq}} < \varepsilon$. Entonces, si $n > \text{máx}(N_1, N_2, N_3)$

$$\left|P_n[a,b) - \int_a^b \Pi_n(x)dx\right| < 2\sqrt{\frac{2}{\pi npq}} < \varepsilon.$$

■

Lema 15 *Sean $a \leqq b$ reales. Entonces*

$$\begin{array}{rcl} \Pi_n(x)dx & \to & \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2_{(n,m)}}{2}} \\ n & \to & \infty \end{array}$$

uniformemente en el intervalo $[a, b]$.

Demostración. Sea $\varepsilon > 0$. Sea N_1 tal que si $n > N_1$ entonces

$$\begin{array}{rcl} x_{(n,0)} & < & a-1 \\ b+1 & < & x_{(n,n+1)}. \end{array}$$

Sea N_2 tal que $n > N_2$ entonces para todo $m = 0, 1, ..., n$ tal que si $a - 1 \leqq x_{(n,m)} \leqq b+1$, se tiene que

$$\left| \frac{\sqrt{npq} P_n(m)}{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2_{(n,m)}}{2}}} - 1 \right| < \frac{\varepsilon\sqrt{2\pi}}{2},$$

entonces

$$\left| \sqrt{npq} P_n(m) - \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2_{(n,m)}}{2}} \right| < \frac{\varepsilon}{2} e^{-\frac{x^2_{(n,m)}}{2}} \leqq \frac{\varepsilon}{2}.$$

La función $\Pi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2_{(n,m)}}{2}}$ es continua en $[a-1, b+1]$ y, por tanto, uniformemente continua.

Sea $\delta > 0$ tal que $\delta < 1$ y $|x - y| < \delta$ entonces $|\Pi(x) - \Pi(y)| < \frac{\varepsilon}{2}$ para $x, y \in [a-1, b+1]$.

Sea N_3 tal que $n > N_3$ entonces $\frac{1}{\sqrt{npq}} < \delta$.

Se define $N = \text{máx}(N_1, N_2, N_3)$. Si $n > N$ y $x \in [a, b]$. Sea m tal que $a - 1 \leqq x_{(n,m)} \leqq b + 1$ y $x_{(n,m)} \leqq x \leqq x_{(n,m+1)}.$ Entonces

$$\begin{array}{rcl} |\Pi(x) - \Pi_n(x)| & = & \left|\Pi(x) - \Pi(x_{(n,m)}) + \Pi(x_{(n,m)}) - \Pi_n(x_{(n,m)}) + \Pi_n(x_{(n,m)}) - \Pi_n(x)\right| \\ & \leqq & \left|\Pi(x) - \Pi(x_{(n,m)})\right| + \left|\Pi(x_{(n,m)}) - \Pi_n(x_{(n,m)})\right| + \left|\Pi_n(x_{(n,m)}) - \Pi_n(x)\right| \\ & < & \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{array}$$

■

Corolario 16 *Sean $A \leqq B$ reales. Dado $\varepsilon > 0$ existe N tal que $n > N$ y $A \leqq a \leqq b \leqq B$ entonces*

$$\left| \int_a^b \Pi_n(x)dx - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \right| < \varepsilon.$$

Demostración. Sea $\varepsilon > 0$, existe N tal que si $n > N$ entonces para todo $x \in [A, B]$, por la convergencia uniforme del lema 15 se cumple que

$$\left| \Pi_n(x)dx - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \right| < \frac{\varepsilon}{B - A}.$$

Si $n > N$ y $A \leqq a \leqq b \leqq B$

$$\begin{aligned}\left|\int_a^b \Pi_n(x)dx - \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx\right| &\leqq \int_a^b \left|\Pi_n(x)dx - \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\right| dx \\ &\leqq \frac{\varepsilon}{B-A}(b-a) \leqq \varepsilon.\end{aligned}$$

■

Corolario 17 *Sean $A \leqq B$ dos reales. Para cada $\varepsilon > 0$ existe N tal que si $n > N$, entonces para todo par (a, b) tal que $A \leqq a \leqq b \leqq B$ se cumple que*

$$\left|P_n\left[a,b\right) - \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx\right| < \varepsilon.$$

Demostración. Sea $\varepsilon > 0$. Existe N_1 tal que si $n > N_1$ y $A \leqq a \leqq b \leqq B$, entonces

$$\left|\int_a^b \Pi_n(x)dx - \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx\right| < \frac{\varepsilon}{2}.$$

Al mismo tiempo, existe N_2 tal que si $n > N_2$ y $A \leqq a \leqq b \leqq B$, se tiene que

$$\left|P_n\left[a,b\right) - \int_a^b \Pi_n(x)dx\right| < \frac{\varepsilon}{2}.$$

Si $n > N = \text{máx}(N_1, N_2)$ y $A \leqq a \leqq b \leqq B$, tenemos que

$$\begin{aligned}\left|P_n\left[a,b\right) - \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\right| &\leqq \left|P_n\left[a,b\right) - \int_a^b \Pi_n(x)dx\right| + \left|\int_a^b \Pi_n(x)dx - \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx\right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

■

Con este corolario se prueba el caso en que a y b son finitos, pero todavía queda pendiente el caso en el cual se elimina esta restricción. Y este caso es precisamente el que se llama teorema integral de DeMoivre-Laplace.

Teorema 18 (Integral de De Moivre-Laplace) *Sean $a \leqq b$ reales, para todo $\varepsilon > 0$ existe un natural N tal que si $n > N$ y $-\infty \leqq a \leqq b \leqq +\infty$, se tiene que*

$$\left|P_n\left[a,b\right) - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx\right| < \varepsilon.$$

Demostración. *Sea $\varepsilon > 0$. Puesto que $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = 1$, existe $A \in R_+$ suficientemente grande para que*

$$\frac{1}{\sqrt{2\pi}}\int_{-A}^{A} e^{-\frac{x^2}{2}}dx > 1 - \frac{\varepsilon}{4}.$$

Entonces, ya que

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-A} e^{-\frac{x^2}{2}}dx + \frac{1}{\sqrt{2\pi}}\int_{-A}^{A} e^{-\frac{z^2}{2}}dx + \frac{1}{\sqrt{2\pi}}\int_{A}^{\infty} e^{-\frac{z^2}{2}}dx = 1,$$

se tiene

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-A} e^{-\frac{x^2}{2}}dx + \frac{1}{\sqrt{2\pi}}\int_{A}^{\infty} e^{-\frac{x^2}{2}}dx < \frac{\varepsilon}{4},$$

por simetría

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-A} e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\int_{-A}^{\infty} e^{-\frac{x^2}{2}}dx < \frac{\varepsilon}{8}.$$

Aplicando el corolario 17 a $-A < A$, *se tiene que existe* N_1 *tal que* $n > N_1$ *entonces*

$$\frac{1}{\sqrt{2\pi}}\int_{-A}^{A} e^{-\frac{x^2}{2}}dx - \frac{\varepsilon}{4} < P_n\left[-A, A\right) < \frac{1}{\sqrt{2\pi}}\int_{-A}^{A} e^{-\frac{x^2}{2}}dx + \frac{\varepsilon}{4}$$

entonces

$$P_n\left[-A, A\right) > 1 - \frac{\varepsilon}{4} - \frac{\varepsilon}{4} = 1 - \frac{\varepsilon}{2},$$

de aquí que

$$P_n\left[-\infty, -A\right) + P_n\left[A, +\infty\right) = 1 - P_n\left[-A, A\right) < \frac{\varepsilon}{2}.$$

Se debe demostrar ahora que, para cada a *y* b $(-\infty \leqq a \leqq b \leqq +\infty)$,

$$\left|P_n\left(a, b\right) - \frac{1}{\sqrt{2\pi}}\int_{a}^{b} e^{-\frac{x^2}{2}}dx\right| < \varepsilon$$

y así el teorema estaría completo.

Para hacer esto, es necesario considerar cada uno de las diferentes formas en las que los puntos a *y* b *pueden ser situados con respecto al intervalo* $(-A, A)$. *Por ejemplo, se examina el caso en que* $a \leqq -A < A \leqq b$.

Entonces

$$\frac{1}{\sqrt{2\pi}}\int_{a}^{b} e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\left[\int\limits_{a}^{-A} + \int\limits_{-A}^{A} + \int\limits_{A}^{b} e^{-\frac{x^2}{2}}dx\right]$$

y

$$P_n(a, b) = P_n(a, -A) + P_n(-A, A) + P_n(A, b).$$

Entonces

$$
\begin{aligned}
\left|P_n(a,b)-\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx\right| &\leqq \left|P_n(a,-A)-\frac{1}{\sqrt{2\pi}}\int_a^{-A} e^{-\frac{x^2}{2}}dx\right| + \\
&\quad + \left|P_n(-A,A)-\frac{1}{\sqrt{2\pi}}\int_{-A}^{A} e^{-\frac{x^2}{2}}dx\right| + \left|P_n(A,b)-\frac{1}{\sqrt{2\pi}}\int_A^{b} e^{-\frac{x^2}{2}}dx\right| \\
&\leqq P_n(-\infty,-A)+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-A} e^{-\frac{x^2}{2}}dx + \\
&\quad + \left|P_n(-A,A)-\frac{1}{\sqrt{2\pi}}\int_{-A}^{A} e^{-\frac{x^2}{2}}dx\right| + P_n(A,\infty)-\frac{1}{\sqrt{2\pi}}\int_A^{\infty} e^{-\frac{x^2}{2}}dx \\
&< \frac{\varepsilon}{2}+\frac{\varepsilon}{4}+\frac{\varepsilon}{8}+\frac{\varepsilon}{8}=\varepsilon.
\end{aligned}
$$

■

La ley de los grandes números

A continuación se presenta una de las aplicaicones más importantes al teorema integral de DeMoivre-Laplace. Se estima lo siguiente, si la cantidad de experimentos n es suficientemente grande, es practicamente seguro que la diferencia entre la frecuencia de éxitos en n experimentos y la probabilidad de éxito de un experimento sea arbitrariamente pequeña. Este hecho fue descubierto por James Bernoulli, y se conoce como teorema de Bernoulli.

El teorema de Bernoulli es la forma más sencilla de un conjunto de resultados relativos a la convergencia de frecuencias o promedios, que se conocen como las "Leyes de los grandes números", y se encuentran entre los más importantes teoremas de la teoría de probabilidad. Es a través de estos toeremas que la teoría se vuelve práctica, de modo que estos teoremas son los responsables del éxito de la aplicación de la teoría de la probabilidad a diversas áreas de la ciencia.

Teorema 19 (de Bernoulli) *Considremos una serie de n experimentos independientes, y sea μ la cantidad de éxitos que ocurren en esta serie de experimentos y p la probabilidad de éxito, en donde $0 < p < 1$. Entonces, para todo $\varepsilon > 0$, sucede que*

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n}-p\right|<\varepsilon\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = 1.$$

Demostración.

$$P\left\{\left|\frac{\mu}{n}-p\right|<\varepsilon\right\} = P\left\{-\varepsilon\sqrt{\frac{n}{pq}} < \frac{\mu-np}{\sqrt{npq}} < \varepsilon\sqrt{\frac{n}{pq}}\right\}$$

y esto implica, por el Teorema Integral de DeMoivre-Laplace, que

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu}{n}-p\right|<\varepsilon\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = 1.$$

Así, para toda constante positiva ε, la probabilidad de $\left|\frac{\mu}{n}-p\right|<\varepsilon$ se aproxima a 1. ■

Capítulo 4

Variables aleatorias

Un acercamiento intuitivo a las variables aleatorias

Uno de los conceptos fundamentales en la teoría de la probabilidad es, el de variable aleatoria. Antes de proceder a la definición formal de este concepto, consideremos algunos ejemplos:

Ejemplos

1. El número de partículas cósmicas, que chocan en una porción específica de la superficie de la tierra en un intervalo de tiempo, está sujeto a fluctuaciones que dependen de factores probabilisticos.

2. El número de llamadas recibidas de suscriptores en una central telefónica en un intervalo de tiempo, es también una variable aleatoria.

3. La velocidad de una molécula de gas no permanece constante pero los cambios dependen de su coalición con otras moléculas. Existen muchas coaliciones, incluso en un intervalo corto de tiempo. La variación de la velocidad de una molécula es en sí de carácter aleatorio.

4. En una muestra tomada al azar de 10 países del continente Americano 8 países tienen un ingreso per capita anual de menos de 10,000 dólares y más de 5,000 dólares. [1] En general, las muestras son consideradas como variables aleatorias.

Los ejemplos arriba mencionados nos muestran claramente que las variables aleatorias tienen que ver con los más variados campos de la ciencia. Y aunque estos ejemplos son completamente diferentes, desde el punto de vista matemático todos presentan las mismas características. Es decir,

[1] The World Factbook. Rank Order 2004. www.odci.gov

en cada ejemplo tratamos con una variable que caracteriza al fenómeno en cuestión. Cada una de estas variables es capaz de tomar varios valores, bajo la influencia de la probabilidad

Es imposible especificar el valor que tomará la variable, pues cambia de una manera aleatoria de experimento a experimento. Entonces para conocer una variable aleatoria, es necesario saber que valores puede tomar, aunque esto es todavía insuficiente para hacer deducciones importantes. De hecho en el ejemplo de la velocidad de una molécula, si consideramos el gas a diferentes temperaturas, es posible que la velocidad molecular permanezca igual, aunque los estados del gas sean diferentes.

Así que para asignar una variable aleatoria, no sólo necesitamos los valores que puede tomar si no que tan seguido lo hace, es decir con que probabilidad toma estos valores.

Por lo tanto una variable aleatoria nos servirá para describir eventos y una función de distribución servirá para asignar probabilidades a ciertos eventos definidos en términos de variables aleatorias.

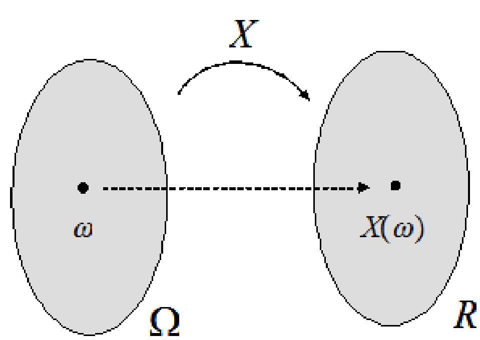
Definición formal

En la primera parte se definió el concepto de espacio de probabilidad que se denotó con la pareja $[\Omega, p]$ donde Ω son todos los posibles resultados de un experimento, y donde p es una función $p : \Omega \longrightarrow R_+$. Con ello podemos proceder a la definición formal de una variable aleatoria.

Definición 10 *Para un espacio de probabilidad* $[\Omega, p]$, *una variable aleatoria denotada por* X *es simplemente una función:*

$$X : \Omega \to R$$

Graficamente una variable aleatoria puede representarse como se muestra en la figura:



El número de valores que puede tomar una variable aleatoria, puede ser discreto o continuo según sea el rango de esta función.

Observación 20 *Una variable aleatoria* X *será discreta, si* X *toma un número finito o infinito numerable de valores reales.*

Por lo pronto nosotros enfocaremos nuestro trabajo únicamente cuando estas variables aleatorias son discretas y en particular finitas.

Ejemplos

A continuación se muestran más ejemplos de variables aleatorias discretas.

1. En el experimento consistente en lanzar dos monedas, el espacio muestral es:

$$\Omega = \{(a,a),(a,s),(s,a),(s,s)\},$$

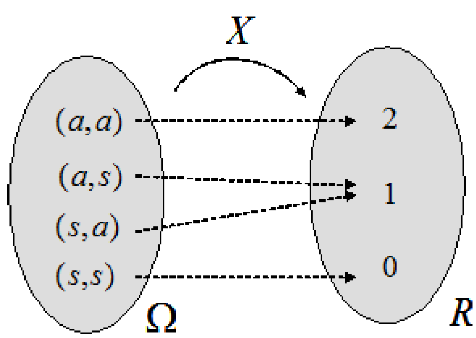
sobre este espacio muestral podemos definir la función:

$$\begin{aligned} X &: \Omega \to R \\ X(\omega) &= \text{número de águilas que tiene cada elemento } \omega. \end{aligned}$$

esta función es una variable aleatoria discreta, ya que toma un número finito de valores reales, a saber

$$\begin{aligned} X(a,a) &= 2 \\ X(a,s) &= X(s,a) = 1 \\ X(s,s) &= 0, \end{aligned}$$

y su representación gráfica es:



2. En el experimento del dado sea $\Omega = \{1,2,3,4,5,6\}$. Si la variable aleatoria está definida como

$$X(\omega) = 10i \text{ si } \omega = i$$

entonces se tiene que la v.a. toma los siguientes valores

$$\begin{aligned} X(1) &= 10,\ X(2) = 20,\ X(3) = 30, \\ X(4) &= 40,\ X(5) = 50 \text{ y } X(6) = 60. \end{aligned}$$

3. Al lanzar una moneda, se tiene que $\Omega = \{aguila, sol\}$.

$$\begin{aligned} \text{Sea } X &: \Omega \to R \\ X(\omega) &= \begin{cases} 1 \text{ si } \omega \text{ es aguila} \\ 0 \text{ si } \omega \text{ es sol} \end{cases}. \end{aligned}$$

4. Al lanzar un par de dados, se tiene que

$$\Omega = \left\{(i,j) \text{ tal que } i = \overline{1,6} \;\text{ y }\; j = \overline{1,6}\;\right\},$$

y se pueden definir diversas variables aleatorias por ejemplo

$$\begin{aligned} \text{Sea } X &: \;\Omega \to R \\ X(\omega) &= \;\text{la suma de las caras de ambos dados} \\ X(\omega) &= \;i + j \;\text{ si }\; \omega = (i,j) \end{aligned}$$

$$\begin{aligned} \text{Sea } Y &: \;\Omega \to R \\ Y(\omega) &= \;\text{el valor absoluto de la diferencia } i - j \\ Y(\omega) &= \;|i - j| \;\text{ si }\; \omega = (i,j). \end{aligned}$$

Hasta ahora hemos definido una variable aleatoria y hemos presentado algunos ejemplos. Pero, sería bueno hablar también de la probabilidad con la que la variable aleatoria X toma el valor x_i. A continuación se hace referencia a la notación que usaremos y al cálculo de dicha probabilidad.

Es necesario saber que para toda i, el conjunto

$$\{\omega \in \Omega : X(\omega) = x_i\}$$

es un evento, donde x_i es el valor que puede tomar la variable aleatoria. Para denotar la probabilidad con la que sucede ese evento lo hacemos con

$$P\left[\{\omega \in \Omega : X(\omega) = x_i\}\right],$$

y para simplificar la notación escribimos $P[X = x_i]$, que se refiere a la probabilidad del conjunto de elementos de Ω que bajo X van a dar a x_i.

Un caso particular, lo podemos retomar del ejemplo 1 de la página 67 en donde la variable aleatoria toma los valores

$$\begin{aligned} X(a,a) &= 2 \\ X(a,s) &= X(s,a) = 1 \\ X(s,s) &= 0. \end{aligned}$$

Se quisiera obtener $P[X = 2]$ que corresponde a la probabilidad del conjunto de elementos de Ω que bajo X van a dar al 2. Análogamente para $P[X = 1]$ y $P[X = 0]$.

Sólo (a,a) elemento de Ω va a dar a 2 y debido a que el espacio muestra tiene cuatro elementos $P[X = 2] = \frac{1}{4}$. Para el caso de $P[X = 1]$ se tiene que bajo la función X, (a,s) y (s,a) van a dar a 1 por lo tanto

$$P[X = 1] = \frac{2}{4} = \frac{1}{2}.$$

De la misma forma sólo (s,s) van a dar al cero por lo que $P[X = 0] = \frac{1}{4}$.

Algunas veces es necesario calcular la probabilidad de un evento

$$\{\omega \in \Omega : X(\omega) \in A\},$$

donde A es algún subconjunto de los reales en vez de un punto como x_i. Si A es un subconjunto de R y X es una variable aleatoria , entonces

$$\{\omega : X(\omega) \in A\} = \bigcup_{x_i \in A} \{\omega : X(\omega) = x_i\}$$

y donde

$$\bigcup_{x_i \in A} \{\omega : X(\omega) = x_i\}$$

significa la unión de todas las i tales que x_i pertenece A.

Usualmente también se abrevia como $\{X \in A\}$, y su probabilidad se denota como $P[X \in A]$. Si A es un intervalo con extremos a y b, es decir $A = (a, b)$, entonces se acostumbra escribir como $P[a < X < b]$, que significa la probabilidad de que X pueda tomar un valor en el intervalo (a, b). Una notación similar la podemos usar para otro tipo de intervalos, veamos más ejemplos en la siguiente tabla:

Inrervalo A	Probabilidad de que X tome
	un valor en el intervalo A.
$[a, b]$	$P[a \leqq X \leqq b]$
$[a, b)$	$P[a \leqq X < b]$
$(a, b]$	$P[a < X \leqq b]$
(a, b)	$P[a < X < b]$
$(-\infty, b]$	$P[X \leqq b]$

La probabilidad de que X se encuentre dentro del intervalo (a, b) se calcula a partir de la siguiente condición

$$P[a < X < b] = P[X < b] - P[X < a]$$

y para calcular

$$P[X < b]\, y P[X < a]$$

introducimos el concepto de función de distribución que abordaremos en la siguiente sección.

Función de distribución de una variable aleatoria

Continuamos con el objetivo de asignar probabilidades a los valores de las variables aleatorias y queremos asignarlos de manera consistente, por tal motivo se introduce en la teoría de probabilidad el concepto de función de distribución de una variable aleatoria.

Definición 11 *Una función de distribución de una variable aleatoria X, denotada por F_X es una función tal que $F_X : R \to [0, 1]$ en donde para todo número real x se tiene*

$$F_X(x) = P[\{\omega \in \Omega : X(\omega) < x\}]$$

que en forma compacta se puede escribir como

$$F_X = P\left[X < x\right],$$

y que se lee como "la probabilidad de que X pueda tomar un valor menor a x".

Y para ilustrarlo mejor retomemos el experimento consistente en lanzar dos monedas.

Ejemplo 19 *Al lanzar un par de monedas, el espacio muestral es:*

$$\Omega = \left\{(a,a),(a,s),(s,a),(s,s)\right\}.$$

y se definió a la variable aleatoria $X(\omega)$ =número de águilas que aparecen, es decir:

$$X(\omega) = \begin{cases} 0 \ \ si\ \omega = (s,s) \\ 1 \ \ si\ \omega = (a,s)\ \ o\ \omega = (s,a) \\ 2 \ \ si\ \omega = (a,a) \end{cases}$$

las probabilidades con las que toma estos valores suponiendo que las monedas son honestas serán, por lo tanto

$$\begin{aligned} P\left[X=0\right] &= \frac{1}{4} \\ P\left[X=1\right] &= \frac{1}{2} \\ P\left[X=2\right] &= \frac{1}{4} \end{aligned}$$

cumpliendo que

$$\sum_{i=1}^{3} P\left[X = x_i\right] = P\left[X=0\right] + P\left[X=1\right] + P\left[X=2\right] = 1.$$

Por otro lado para construír su función de distribución basta con seguir la definición, es decir encontrar:

$$F_X = P\left[X < x\right]$$

y para facilitar los cálculos, lo podemos ver de la siguiente forma

$$P\left[X < x\right] = \sum_{x_i < x} P\left[X = x_i\right]$$

donde x_i es un valor particular de X.

En nuestro ejemplo X puede tomar los valores $0, 1$ y 2. Si $x \leq 0$, entonces

$$P\left[X < x\right] = \sum_{x_i < x} P\left[X = x_i\right] = 0.$$

Si $0 < x \leq 1$, entonces

$$P\left[X < x\right] = \sum_{x_i < x} P\left[X = x_i\right] = P\left[X = 0\right] = \frac{1}{4}$$

Si $1 < x \leq 2$, entonces

$$\begin{aligned} P[X < x] &= \sum_{x_i < x} P[X = x_i] \\ &= P[X = 0] + P[X = 1] \\ &= \frac{1}{4} + \frac{1}{2} = \frac{3}{4} \end{aligned}$$

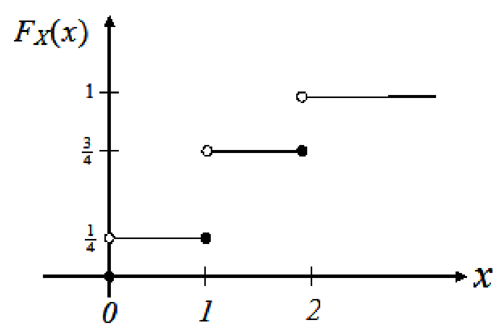
Si $x > 2$, entonces

$$\begin{aligned} P[X < x] &= \sum_{x_i < x} P[X = x_i] \\ &= P[X = 0] + P[X = 1] + P[X = 2] \\ &= \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1 \end{aligned}$$

que en forma compacta lo podemos escribir como

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{4} & \text{si } 0 < x \leq 1 \\ \frac{3}{4} & \text{si } 1 < x \leq 2 \\ 1 & \text{si } x > 2 \end{cases}.$$

Gráficamente, esta función se puede ver como sigue:



Ejemplo 20 *Al lanzar un par de dados, el espacio muestra es el conocido conjunto:*

$$\Omega = \{(i, j) | i = 1, .., 6 \quad y \quad j = 1, .., 6\}.$$

Diversas variables aleatorias se pueden definir, por ejemplo tomar el valor absoluto de la resta de ambos dados es decir:

$$Y(\omega) = |i - j|, \ \ si\ \omega = (i, j),$$

entonces los valores que puede tomar la variable aleatoria son $0, 1, 2, 3, 4, 5$*. Es decir* $Y : \Omega \to \{0, 1, 2, 3, 4, 5\}$ *entonces:*

$$Y(\omega)=\begin{cases} 0 \text{ si } \omega & \in \{(1,1),\ (2,2),\ (3,3),\ (4,4),\ (5,5),\ (6,6)\} \\ 1 \text{ si } \omega & \in \left\{ \begin{array}{c} (2,1),\ (1,2),\ (3,2),\ (2,3),\ (4,3), \\ (3,4)\ (4,5),\ (5,4),\ (6,5),\ (5,6) \end{array} \right\} \\ 2 \text{ si } \omega & \in \{(3,1),\ (1,3),\ (4,2),\ (2,4),\ (5,3),\ (3,5)\ (4,6),\ (6,4)\} \\ 3 \text{ si } \omega & \in \{(4,1),\ (1,4),\ (5,2),\ (2,5),\ (6,3),\ (3,6)\ \} \\ 4 \text{ si } \omega & \in \{(5,1),\ (1,5),\ (6,2),\ (2,6)\} \\ 5 \text{ si } \omega & \in \{\ (6,1),\ (1,6)\} \end{cases}$$

y las probabilidades con las que toma estos valores suponiendo que los dados son honestos son:

$$\begin{aligned} P[Y=0] &= \frac{6}{36}=\frac{1}{6} \\ P[Y=1] &= \frac{10}{36}=\frac{5}{18} \\ P[Y=2] &= \frac{8}{36}=\frac{2}{9} \\ P[Y=3] &= \frac{6}{36}=\frac{1}{6} \\ P[Y=4] &= \frac{4}{36}=\frac{1}{9} \\ P[Y=5] &= \frac{2}{36}=\frac{1}{18}, \end{aligned}$$

en donde es fácil ver que todas ellas suman uno.

Y puede tomar los valores $0,1,2,3,4,5$.

Si $y\le 0$, entonces

$$P[Y<y]=\sum_{x_i<x} P[Y=y_i]=0.$$

Si $0<y\le 1$, entonces

$$\begin{aligned} P[Y<y] &= \sum_{x_i<x} P[Y=y_i] \\ &= P[Y=0]=\frac{1}{6}. \end{aligned}$$

Si $1<y\le 2$, entonces

$$\begin{aligned} P[Y<y] &= \sum_{x_i<x} P[Y=y_i] \\ &= P[Y=0]+P[Y=1] \\ &= \frac{1}{6}+\frac{5}{18}=\frac{4}{9}. \end{aligned}$$

Si $2<y\le 3$, entonces

$$\begin{aligned} P[Y<y] &= \sum_{x_i<x} P[Y=y_i] \\ &= P[Y=0]+P[Y=1]+P[Y=2] \\ &= \frac{1}{6}+\frac{5}{18}+\frac{2}{9}=\frac{2}{3}. \end{aligned}$$

Si $3 < y \leq 4$, entonces

$$\begin{aligned} P[Y < y] &= \sum_{x_i < x} P[Y = y_i] \\ &= P[Y = 0] + P[Y = 1] + P[Y = 2] + P[Y = 3] \\ &= \frac{1}{6} + \frac{5}{18} + \frac{2}{9} + \frac{1}{6} = \frac{5}{6}. \end{aligned}$$

Si $4 < y \leq 5$, entonces

$$\begin{aligned} P[Y < y] &= \sum_{x_i < x} P[Y = y_i] \\ &= P[Y = 0] + P[Y = 1] + P[Y = 2] + P[Y = 3] + P[Y = 4] \\ &= \frac{1}{6} + \frac{5}{18} + \frac{2}{9} + \frac{1}{6} + \frac{1}{9} = \frac{17}{18}. \end{aligned}$$

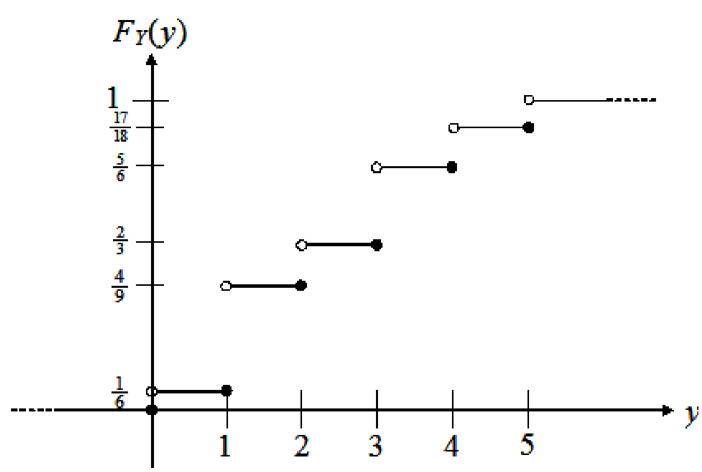
Si $y > 5$, entonces

$$\begin{aligned} P[Y < y] &= \sum_{x_i < x} P[Y = y_i] \\ &= \sum_{i=0}^{5} P[Y = y_i] = 1 \end{aligned}$$

que en forma compacta lo podemos escribir como

$$F_Y(y) = \begin{cases} 0 & \text{si } y \leq 0 \\ \frac{1}{6} & \text{si } 0 < y \leq 1 \\ \frac{4}{9} & \text{si } 1 < y \leq 2 \\ \frac{2}{3} & \text{si } 2 < y \leq 3 \\ \frac{5}{6} & \text{si } 3 < y \leq 4 \\ \frac{17}{18} & \text{si } 4 < y \leq 5 \\ 1 & \text{si } y > 5 \end{cases}.$$

Gráficamente, esta función se puede ver como sigue:



Propiedades de la función de distribución

En esta sección analizaremos las propiedades de la función de distribución, como se ha podido ver en los ejemplos anteriores adopta características muy particulares como la forma de escalera con saltos en los valores aislados que toma la variable aleateatoria, siendo en cada uno de ellos continua por la izquierda. Resumiremos entonces sus características en las siguientes tres propiedades.

1. Una función de disribución es una función monótona no decreciente.
2. $\lim_{n\to-\infty} F_X(n) = 0$ y $\lim_{n\to\infty} F_X(n) = 1$.
3. Una función de distribución es continua por la izquierda.

Proposición 9 *La función de distribución de una variable aleatoria es una función monótona no decreciente.*

Demostración. Si F es una función monótona no decreciente entonces $F_X(x_1) \leqq F_X(x_2)$ para cualquier $x_1 < x_2$. Para demostrarlo determinaremos la probabilidad de la siguiente desigualdad $x_1 \leqq X < x_2$ utilizando la definición de función de distribución de una variable aleatoria.

Sean

$$\begin{aligned} A &= [\omega \in \Omega : X(\omega) < x_2] = [X < x_2] \\ B &= [\omega \in \Omega : X(\omega) < x_1] = [X < x_1] \\ C &= [\omega \in \Omega : x_1 \leqq X(\omega) < x_2] = [x_1 \leqq X < x_2] \end{aligned}$$

entonces

$$[X < x_2] = [X < x_1] \cup [x_1 \leqq X < x_2]$$

obsérvese que $[X < x_1] \cap [x_1 \leqq X < x_2] = \varnothing$ por lo tanto al obtener la probabilidad de la expresión anterior

$$\begin{aligned} P[X < x_2] &= P[[X < x_1] \cup [x_1 \leqq X < x_2]] \\ &= P[X < x_1] + P[x_1 \leqq X < x_2] \end{aligned}$$

entonces

$$P[x_1 \leqq X < x_2] = P[X < x_2] - P[X < x_1]$$

de donde

$$P[x_1 \leqq X < x_2] = F_X(x_2) - F_X(x_1)$$

y por definición sabemos que la probabilidad es un número no negativo, por lo tanto se sigue que

$$F_X(x_2) \geqq F_X(x_1) \text{ para todo } x_2 > x_1.$$

Por lo tanto la función de distribución de toda variable aleatoria es una función no decreciente.

■

A continuación, se demuestra un teorema que servirá para demostrar la proposición que sigue.

Teorema 21 **(i)** *$A_1 \subset A_2 \subset ... \subset$ y definimos $A = \bigcup_{i=1}^{\infty} A_n$, entonces*

$$\lim_{n\to\infty} P(A_n) = P(A)$$

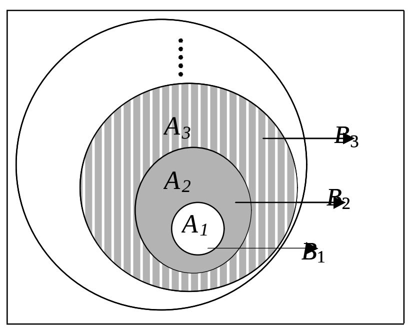
(ii) *$A_1 \supset A_2 \supset ... \supset$ y definimos $A = \bigcap_{i=1}^{\infty} A_n$, entonces también se cumple*

$$\lim_{n\to\infty} P(A_n) = P(A).$$

Demostración. i) Sean los conjuntos:

$B_1 = A_1$
$B_2 = A_2 \cap A_1^c$
$B_3 = A_3 \cap A_2^c$.
$\vdots$
$B_n = A_n \cap A_{n-1}^c$

Por definición los conjuntos B_i son ajenos por parejas, es decir los elementos están en A_i pero no en A_{i-1}, ver la figura siguiente:



entonces

$$A_n = \bigcup_{i=1}^{n} B_i$$

y

$$A = \bigcup_{i=1}^{\infty} B_i$$

en consecuencia,

$$P(A_n) = \sum_{i=1}^{n} P(B_i)$$

y

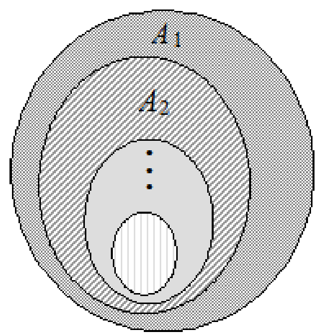
$$P(A) = \sum_{i=1}^{\infty} P(B_i).$$

Entonces

$$\begin{aligned}\lim_{n\to\infty} F_X(n) &= \lim_{n\to\infty} P\left[X < n\right] \\ &= \lim_{n\to\infty} P(A_n) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) \\ &= \sum_{i=1}^{\infty} P(B_i) \\ &= P(A).\end{aligned}$$

■

Demostración. ii) Sea $A_1 \supset A_2 \supset ... \supset$, ver la figura siguiente:



de donde se deduce que $A = \bigcap_{i=1}^{\infty} A_n$ entonces $A_1^c \subset A_2^c \subset ...$ y por la Ley de De Morgan se tiene que:

$$A^c = (\bigcap_{i=1}^{\infty} A_n)^c = (\bigcup_{i=1}^{\infty} A_n^c)$$

entonces por el inciso anterior

$$\lim_{n\to\infty} P(A_n^c) = P(A^c)$$

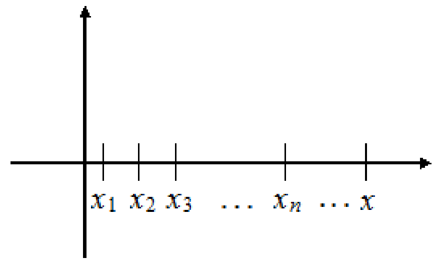
como $P(A_n) = 1 - P(A_n^c)$ y $P(A) = 1 - P(A^c)$, entonces

$$\begin{aligned}\lim_{n\to\infty} P(A_n) &= \lim_{n\to\infty} [1 - P(A_n^c)] \\ &= 1 - \lim_{n\to\infty} P(A_n^c) \\ &= 1 - P(A^c) \\ &= 1 - [1 - P(A)] \\ &= P(A).\end{aligned}$$

■

Proposición 10

(i) $\lim_{n\to\infty} F_X(n) = 1$ y



(ii) $\lim\limits_{n\to-\infty} F_X(n) = 0.$

Demostración. **(i)** Se construyen los siguientes conjuntos

Sea $B_n = \{\omega \in \Omega : X(\omega) < n\}$ de donde $... \subseteq B_{-2} \subseteq B_{-1} \subseteq B_0 \subseteq B_1 \subseteq B_2 ...$ y se tiene que

$$\bigcup_{i=0}^{\infty} B_n = \Omega$$

y

$$\bigcap_{i=0}^{-\infty} B_n = \varnothing$$

entonces por el teorema 21 de la página 75

$$\lim_{n\to\infty} F_X(n) = \lim_{n\to\infty} P\left[X < n\right] = \lim_{n\to\infty} P(B_n) = P(\Omega) = 1,$$

análogamente

$$\lim_{n\to-\infty} F_X(n) = \lim_{n\to-\infty} P\left[X < n\right] = \lim_{n\to-\infty} P(B_n) = P(\varnothing) = 0.$$

■

Proposición 11 *Las funciones de distribución son continuas por la izquierda.*

Demostración. Sea $x_0 < x_1 < x_2 < ... < x_n ...$ una sucesión creciente que converge a x.

Sea A_n el evento formado por

$$A_n = \{x_n \leqq X < x\}.$$

Por ejemplo en el caso de A_1 y A_2, se tiene que

$$\begin{aligned} A_1 &= \{x_1 \leqq X < x\} \\ A_2 &= \{x_2 \leqq X < x\} \end{aligned}$$

entonces $A_2 \subset A_1$.

Es claro que $A_i \subset A_j$ siempre que $i > j$ y además la intersección de todos estos eventos es el vacío $\bigcap_{i=1}^{\infty} A_n = \varnothing$, entonces por la proposición 10 se tiene que:

$$\begin{aligned} F_X(x) - F_X(x-0) &= F_X(x) - \lim_{n\to\infty} F_X(x_n) \\ &= \lim_{n\to\infty} \left(F_X(x) - F_X(x_n)\right) \\ &= \lim_{n\to\infty} \left(P\left[X < x\right] - P\left[X < x_n\right]\right) \\ &= \lim_{n\to\infty} P\left[x_n \leqq X < x\right] \\ &= \lim_{n\to\infty} P\left[A_n\right] \\ &= P(\varnothing) = 0. \end{aligned}$$

Por lo tanto se demuestra la continuidad por la izquierda $F_X(x) = F_X(x-0)$. ∎

Características numéricas de las variables aleatorias

Las características numéricas de las variables aleatorias, serán constantes que sirven para dar una caracterización cuantitativa general de las variables aleatorias.

La esperanza y la varianza son particularmente importantes.

Esperanza

Empezaremos la discusión considerando el siguiente ejemplo ilustrativo. Se realiza un juego que consiste en lanzar un dado honesto, el pago que recibe el jugador será de 10 pesos multiplicados por la cara del dado que se ha obtenido y para poder jugar se deben pagar r pesos por juego. La pregunta es ¿en qué condiciones conviene jugar este juego?

Si el juego se realiza un número grande de veces, digamos 6 millones de veces entonces se deben pagar para poder jugar $r(6,000,000)$ y como el dado es honesto entonces cada cara caerá aproximadamente $1,000,000$ de veces por lo tanto el pago que esperaríamos recibir es el siguiente:

$$10(10^6) + 20(10^6) + 30(10^6) + 40(10^6) + 50(10^6) + 60(10^6)$$

por lo que el pago promedio recibido es:

$$\begin{aligned} &\frac{10(10^6) + 20(10^6) + 30(10^6) + 40(10^6) + 50(10^6) + 60(10^6)}{6 * 10^6} \\ =\;& 10\left(\frac{1}{6}\right) + 20\left(\frac{1}{6}\right) + 30\left(\frac{1}{6}\right) + 40\left(\frac{1}{6}\right) + 50\left(\frac{1}{6}\right) + 60\left(\frac{1}{6}\right) \\ =\;& \left(\frac{1}{6}\right)(210) \\ =\;& 35 \end{aligned}$$

y esta cantidad corresponde al valor esperado de ganancia si se juega una sola vez.

Para decidir si nos conviene jugar tomamos en cuenta lo siguiente.

$$\text{Si el } \textit{"valor esperado"} \begin{cases} \text{es mayor que } r \text{ habrá una ganancia} \\ \text{es menor que } r \text{ habrá una pérdida} \\ \text{es igual que } r \text{ no ganas ni pierdes} \end{cases}$$

Cuestiones análogas, que involucran el cálculo del valor en promedio de una variable aleatoria, se encuentran en los más variados tipos de problemas. A continuación se define la esperanza de una variable aleatoria.

Definición 12 *Sean $X(a_1), X(a_2), X(a_3), ..., X(a_n)$ los posibles valores que puede tomar una variable aleatoria X y sean*

$$P(a_1), P(a_2), P(a_3), ..., P(a_n)$$

sus probabilidades correspondientes. Se define a la esperanza, de la variable aleatoria X, como la suma $\sum_{i=1}^{n} X(a_i)\, P(a_i)$ y se denota con $E(X)$.

Teoremas sobre la esperanza

Teorema 22 *La esperanza de una constante es una constante.*

$$E(C) = C$$

Demostración. Sea C una constante y $X : \Omega \longrightarrow R$ la variable aleatoria tal que $X = C$ entonces $X(a_i) = C$ para todo $a_i \in \Omega$.

$$\begin{aligned} E(X) &= \sum_{i=1}^{n} X(a_i)\, P(a_i) \\ &= \sum_{i=1}^{n} CP(a_i) \\ &= C\sum_{i=1}^{n} P(a_i) \\ &= C(1) \\ &= C \end{aligned}$$

■

Teorema 23 *La esperanza de una constante por una variable aleatoria es igual a la constante por la esperanza de la variable aleatoria.*

$$E(CX) = CE(X)$$

Demostración. Sea C una constante y sean $X : \Omega \longrightarrow R$ y $Z : \Omega \longrightarrow R$ variables aleatorias tal que $Z = CX$ entonces $Z(a_i) = CX(a_i)$ para todo $a_i \in \Omega$.

$$\begin{aligned}
E(Z) &= \sum_{i=1}^{n} Z(a_i) P(a_i) \\
&= \sum_{i=1}^{n} CX(a_i) P(a_i) \\
&= C \sum_{i=1}^{n} X(a_i) P(a_i) \\
&= CE(X)
\end{aligned}$$

■

Teorema 24 *La esperanza de la suma de dos variables aleatorias es igual a la suma de sus esperanzas.*

$$E(X+Y) = E(X) + E(Y)$$

Demostración. Sean $X, Y, Z : \Omega \longrightarrow R$ variables aleatorias tal que $Z = X + Y$ entonces $Z(a_i) = X(a_i) + Y(a_i)$ para todo $a_i \in \Omega$.

$$\begin{aligned}
E(Z) &= \sum_{i=1}^{n} Z(a_i) P(a_i) \\
&= \sum_{i=1}^{n} [X(a_i) + Y(a_i)] P(a_i) \\
&= \sum_{i=1}^{n} X(a_i) P(a_i) + Y(a_i) P(a_i) \\
&= \sum_{i=1}^{n} X(a_i) P(a_i) + \sum_{i=1}^{n} Y(a_i) P(a_i) \\
&= E(X) + E(Y)
\end{aligned}$$

■

Teorema 25 *La esperanza de la suma de un número finito de variables aleatorias es igual a la suma de sus esperanzas.*

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n).$$

Demostración. Sean $X_1, X_2, ..., X_n, Z : \Omega \longrightarrow R$, variables aleatorias, $Z = X_1 + X_2 + ... + X_n$ y análogamente al teorema que se acaba de demostrar se tiene que

$$\begin{aligned}
E(Z) &= \sum_{i=1}^{n} Z\,(a_i)\,P(a_i) \\
&= \sum_{i=1}^{n} [X_1(a_i) + X_2(a_i) + ... + X_n(a_i)]\,P(a_i) \\
&= \sum_{i=1}^{n} X_1(a_i)P(a_i) + \sum_{i=1}^{n} X_2(a_i)P(a_i) + ... + \sum_{i=1}^{n} X_n(a_i)P(a_i) \\
&= E(X_1) + E(X_2) + ... + E(X_n)
\end{aligned}$$

■

Ejemplo 21 *Una compañia vende productos para consumo local y para exportación, obtener el número estimado de ventas para el próximo año con los siguientes datos.*

Se venden X unidades para consumo local con las siguientes probabilidades

Solución 4

	$X(a_i)$	$P(a_i)$
a_1	1000	0,1
a_2	3000	0,3
a_3	5000	0,4
a_4	10000	0,2

Se venden Y unidades para exportación con las siguientes probabilidades

Solución 5

	$Y(a_i)$	$P(a_i)$
a_1	300	0,4
a_2	500	0,5
a_3	700	0,1

Entonces el número esperado de ventas locales es

$$\begin{aligned}
E(X) &= \sum_{i=1}^{4} X\,(a_i)\,P(a_i) \\
&= 1000\,(0{,}1) + 3000\,(0{,}3) + 5000\,(0{,}4) + 10000\,(0{,}2) \\
&= 5000
\end{aligned}$$

y el número esperado de exportaciones es

$$\begin{aligned}
E(Y) &= \sum_{i=1}^{3} Y\,(a_i)\,P(a_i) \\
&= 300\,(0{,}4) + 500\,(0{,}5) + 700\,(0{,}1) \\
&= 440
\end{aligned}$$

La compañia tiene una ganancia de \$2000 por cada unidad vendida en el mercado local y de \$3500 en cada unidad exportada. Por tanto la ganancia total es de

$$T = 2000X + 3500Y$$

De tal forma que usando las propiedades de la esperanza se tiene que la ganancia esperada para el siguiente año será de

$$\begin{aligned} E(T) &= E(2000X + 3500Y) \\ &= E(2000X) + E(3500Y) \\ &= 2000E(X) + 3500E(Y) \\ &= 2000 * 5000 + 3500 * 440 \\ &= 11,540,000 \end{aligned}$$

Varianza

Definición 13 *La varianza de una variable aleatoria X está definida como la esperanza del cuadrado de la desviación de X con $E(X)$.*[2]

$$Var(X) = E\left[X - E(X)\right]^2$$

Usando la definición de esperanza y desarrollando el cuadrado se tiene

$$\begin{aligned} Var(X) &= \sum_{i=1}^{n} \left[X(a_i) - E(X)\right]^2 P(a_i) \\ &= \sum_{i=1}^{n} \left[X^2(a_i) - 2X(a_i)E(X) + E^2(X)\right] P(a_i) \\ &= \sum_{i=1}^{n} X^2(a_i)P(a_i) - \sum_{i=1}^{n} 2X(a_i)E(X)P(a_i) + \sum_{i=1}^{n} E^2(X)P(a_i) \\ &= E(X^2) - 2E(X)\sum_{i=1}^{n} X(a_i)P(a_i) + E^2(X)\sum_{i=1}^{n} P(a_i) \\ &= E(X^2) - 2E(X)E(X) + E^2(X) * 1 \\ &= E(X^2) - 2E^2(X) + E^2(X) \\ &= E(X^2) - E^2(X) \end{aligned}$$

por lo tanto la varianza también se puede expresar como

$$Var(X) = E(X^2) - E^2(X) \tag{4.1}$$

[2]La raiz cuadrada de la varianza es llamada desviación estándar y se denota como $\sigma = \sqrt{Var(X)}$

Ejemplo 22 *La variable aleatoria X es "el número de estudiantes mujeres en un salón de clases con 5 alumnos" y las probabilidades estan distribuídas en la siguiente tabla, obtener el número de mujeres esperado en un grupo y la varianza de X.*

	$X(a_i)$	$P(a_i)$
a_1	0	$\frac{1}{32}$
a_2	1	$\frac{5}{32}$
a_3	2	$\frac{10}{32}$
a_4	3	$\frac{10}{32}$
a_5	4	$\frac{5}{32}$
a_6	5	$\frac{1}{32}$

Entonces la esperanza es

$$\begin{aligned} E(X) &= \sum_{i=1}^{6} X\left(a_i\right) P(a_i) \\ &= 0\left(\frac{1}{32}\right)+1\left(\frac{5}{32}\right)+2\left(\frac{10}{32}\right)+3\left(\frac{10}{32}\right)+4\left(\frac{5}{32}\right)+5\left(\frac{1}{32}\right) \\ &= \frac{80}{32}=2{,}5, \end{aligned}$$

es decir, en promedio los grupos con 5 personas tienen 2,5 mujeres.

$$\begin{aligned} E(X^2) &= \sum_{i=1}^{6} X^2\left(a_i\right) P(a_i) \\ &= 0^2\left(\frac{1}{32}\right)+1^2\left(\frac{5}{32}\right)+2^2\left(\frac{10}{32}\right)+3^2\left(\frac{10}{32}\right)+4^2\left(\frac{5}{32}\right)+5^2\left(\frac{1}{32}\right) \\ &= 7{,}5 \end{aligned}$$

Para obtener la varianza

$$\begin{aligned} Var(X) &= E(X^2)-E^2(X) \\ &= 7{,}5-(2{,}5)^2 \\ &= 7{,}5-6{,}25 \\ &= 1{,}25 \end{aligned}$$

Teoremas sobre la varianza

Teorema 26 *La varianza de una constante es cero.*

$$Var(C)=0$$

Demostración. Sea C una constante y $X : \Omega \longrightarrow R$ la variable aleatoria tal que $X = C$, entonces $X(a_i) = C$ para todo $a_i \in \Omega$.

$$\begin{aligned} Var(X) &= E\left[X - E(X)\right]^2 \\ &= \sum_{i=1}^{n} \left[X\left(a_i\right) - E(C)\right]^2 P(a_i) \\ &= \sum_{i=1}^{n} \left[C - C\right]^2 P(a_i) \\ &= 0 \end{aligned}$$

■

Teorema 27 *Si C es una constante, entonces*

$$Var(CX) = C^2 Var(X)$$

Demostración. Sea C una constante y sean $X : \Omega \longrightarrow R$ y $Z : \Omega \longrightarrow R$ variables aleatorias tal que $Z = CX$ entonces $Z(a_i) = CX(a_i)$ para todo $a_i \in \Omega$.

$$\begin{aligned} Var(Z) &= E\left[Z - E(Z)\right]^2 \\ &= \sum_{i=1}^{n} \left[Z\left(a_i\right) - E(CX)\right]^2 P(a_i) \\ &= \sum_{i=1}^{n} \left[CX\left(a_i\right) - CE(X)\right]^2 P(a_i) \\ &= \sum_{i=1}^{n} C^2 \left[X\left(a_i\right) - E(X)\right]^2 P(a_i) \\ &= C^2 \sum_{i=1}^{n} \left[X\left(a_i\right) - E(X)\right]^2 P(a_i) \\ &= C^2 E\left[X - E(X)\right]^2 \\ &= C^2 Var(X). \end{aligned}$$

■

Independencia de variables aleatorias

Uno de los conceptos mas importantes en teoría de la probabilidad es la independencia de eventos y mantiene su importancia para variables aleatorias. Abordaremos la definición de independencia para dos variables aleatorias y posteriormente la generalización a mas variables.

Entonces de acuerdo con la definición de independencia de eventos, consideremos dos variables aleatorias X y Y. Sean x_1, $x_2, ...,$ x_n los valores que toma la variable X y y_1, $y_2, ...,$ y_n los valores que toma la variable Y. Entonces,

Definición 14 *Las variables aleatorias X y Y son independientes si para cualquier x_i y y_j se verifica*

$$P\left[(X = x_i) \cap \ (Y = y_j)\right] = P(X = x_i)P(Y = y_j).$$

Ejemplo 23 *Se lanzan dos monedas honestas*

$\Omega = \{(a,a),(a,s),(s,a),(s,s)\}$
Sea la variable aleatoria $X =$ al número de soles que tiene cada elemento ω.
$X(\omega) = \begin{cases} 0 \\ 1 \\ 2 \end{cases}.$
Las probabilidades con las que puede tomar cada valor son:

$$\begin{aligned} P\left(X = 0\right) &= \frac{1}{4} \\ P\left(X = 1\right) &= \frac{1}{2} \\ P\left(X = 2\right) &= \frac{1}{4}. \end{aligned}$$

Sea Y la variable aleatoria que asigna
$Y(\omega) = \begin{cases} 1 & \text{si sale alguna aguila} \\ 2 & \text{si no sale aguila} \end{cases}.$
Las probabilidades con las que puede tomar cada valor son:

$$\begin{aligned} P\left(Y = 1\right) &= \frac{3}{4} \\ P\left(Y = 2\right) &= \frac{1}{4}. \end{aligned}$$

Para verificar que X y Y son independientes basta con que cumpla la definición, si calculamos la probabilidad de que se cumpla la intersección

$$P\left[(X = 0) \cap \ (Y = 1)\right] = \frac{1}{4},$$

y por otro lado

$$P(X = 0)P(Y = 1) = \left(\frac{1}{4}\right)\left(\frac{3}{4}\right) = \frac{3}{16}.$$

Por tanto

$$P\left[(X = 0) \cap \ (Y = 1)\right] \neq P(X = 0)P(Y = 1)$$

entonces X y Y no son independientes.

En el siguiente ejemplo se considera el experimento de lanzar una moneda y lanzar un dado. Intuitivamente, se cree que cualquiera que sea la cara de la moneda, no deberá tener influencia sobre el resultado del dado, entonces estos eventos deberían ser independientes, pero vamos a analizarlo más detalladamente.

$\Omega = \{a, s\}$
$\Lambda = \{1, 2, 3, 4, 5, 6\}$
$\Omega \times \Lambda = \left\{ \begin{array}{c} (a,1),(a,2),(a,3),(a,4),(a,5),(a,6) \\ (s,1),(s,2),(s,3),(s,4),(s,5),(s,6) \end{array} \right\}$
Sea X la variable aleatoria que asigna
$X(\omega) = \begin{cases} 1 & \text{si sale aguila} \\ 0 & \text{si sale sol} \end{cases}.$
Sea Y la variable aleatoria que asigna
$Y(\omega) = \{ \; i \quad \text{si sale la cara } i \; .$
Entonces $P\left[(X = 1) \cap \; (Y = 1)\right]$ sólo sucede en el evento $(a, 1)$, por lo que su probabilidad es:

$$P\left[(X = 1) \cap \; (Y = 1)\right] = \frac{1}{12}$$

y por otro lado $(X = 1)$ sucede en todos los elementos que tienen águila por lo que su probabilidad es:

$$P(X = 1) = \frac{6}{12} = \frac{1}{2}$$

y $(Y = 1)$ sucede en los eventos que tienen uno por lo que su probabilidad es:

$$P(Y = 1) = \frac{2}{12} = \frac{1}{6}$$

entonces

$$P(X = 1)P(Y = 1) = \frac{1}{12}$$

por lo tanto las variables aleatorias X y Y son independientes.

Contextualización de la independencia entre variables aleatorias en espacios producto

Ahora, propondremos un análisis similar al expuesto en el capítulo de espacios productos de probabilidad, pero considerando variables aleatorias.

Se tienen $E_1 = [\Omega_1, p_1]$ y $E_2 = [\Omega_2, p_2]$ dos espacios de probabilidad finitos, y dos variables aleatorias $X_1 : \Omega_1 \to R$ y $X_2 : \Omega_2 \to R$, las cuales inducen el espacio producto $E = [\Omega, P]$, con $\Omega = \Omega_1 \times \Omega_2$ y con $P(a_1, a_2) = p_1(a_1)p_2(a_2)$. Además considérense las siguientes variables aleatorias que se definen como a continuación:

$$\begin{aligned} \overline{X_1} &: \Omega_1 \times \Omega_2 \to R \\ \overline{X_1}(a_1, a_2) &= X_1(a_1), \end{aligned}$$

y

$$\begin{aligned} \overline{X_2} &: \Omega_1 \times \Omega_2 \to R \\ \overline{X_2}(a_1, a_2) &= X_2(a_2), \end{aligned}$$

entonces $\overline{X_1}$ y $\overline{X_2}$ serán independientes.

Y ahora generalizando la situación anterior se tienen n variables aleatorias

$$\begin{array}{rcl} [\Omega_1, p_1] \text{ con } X_1 & : & \Omega_1 \to R \\ [\Omega_2, p_2] \text{ con } X_2 & : & \Omega_2 \to R \\ & \vdots & \\ [\Omega_n, p_n] \text{ con } X_n & : & \Omega_n \to R \end{array}$$

en donde se define a $[\Omega, P]$ tal que $\Omega = \Omega_1 \times \Omega_2 \times ... \times \Omega_n$ y con P una función $P : \Omega \longrightarrow R_+$ tal que

$$P(a_1, a_2, ..., a_n) = p_1(a_1)p_2(a_2)...p_n(a_n)$$

y además las variables aleatorias

$$\begin{array}{rcl} \overline{X_i} & : & \Omega \to R \\ \overline{X_i}(a_1, a_2, ..., a_n) & = & X_i(a_i) \end{array}$$

para toda $i = 1, ..., n$, ahora todas las variables aleatorias $\overline{X_i}$ serán independientes.

Así tendremos una manera fácil de poder generar un gran número de variables aleatorias que serán independientes.

Ejemplo 24 *El espacio de probabilidad que se genera al lanzar dos dados honestos de diferentes colores es el siguiente: Sea* $[\Omega_1, p_1]$ *con* $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$ *y sea* X_1 *la variable aleatoria, el número de puntos que se obtienen en el dado rojo, entonces*

$$\begin{array}{rcl} X_1 & : & \Omega_1 \to R \\ X_1(\omega) & = & \{i \text{ cuando la cara del dado cae } i\ . \end{array}$$

Y sea $E_2 = [\Omega_2, p_2]$ *el espacio de probabilidad generado al lanzar el dado, es decir con* $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$ *y sea* X_2 *el número de puntos que salen en el dado azul, entonces*

$$\begin{array}{rcl} X_2 & : & \Omega_2 \to R \\ X_2(\omega) & = & \{i \text{ cuando la cara del dado cae } i\ . \end{array}$$

El espacio producto generado es:

$$\Omega_1 \times \Omega_2 = \left\{ \begin{array}{c} (1,1), (1,2), (1,3), (1,4), (1,5)\ , (1,6,), \\ (2,1), (2,2), (2,3), (2,4), (2,5)\ , (2,6,), \\ (3,1), (3,2), (3,3), (3,4), (3,5)\ , (3,6,), \\ (4,1), (4,2), (4,3), (4,4), (4,5)\ , (4,6,), \\ (5,1), (5,2), (5,3), (5,4), (5,5)\ , (5,6,), \\ (6,1), (6,2), (6,3), (6,4), (6,5)\ , (6,6,) \end{array} \right\}$$

y las variables aleatorias $\overline{X_1} : \Omega_1 \times \Omega_2 \to R$ *y* $\overline{X_2} : \Omega_1 \times \Omega_2 \to R$ *serán independientes. Para el caso:*

$$\begin{array}{rcl} \overline{X}_1(1,1) & = & X_1(1) = 1 \\ \overline{X}_1(2,1) & = & X_1(2) = 2 \\ \overline{X}_1(3,1) & = & X_1(3) = 3 \\ \overline{X}_1(4,1) & = & X_1(4) = 4 \end{array}$$

y

$$\begin{aligned}
\overline{X}_2(1,1) &= X_2(1) = 1\\
\overline{X}_2(2,1) &= X_2(1) = 1\\
\overline{X}_2(3,1) &= X_2(1) = 1\\
\overline{X}_2(4,1) &= X_2(1) = 1
\end{aligned}$$

$$\begin{aligned}
P\left[\left(\overline{X}_1(w)=1\right)\cap\left(\overline{X}_2(w)=1\right)\right] &= P\left(\overline{X}_1(w)=1\right)\cdot P\left(\overline{X}_2(w)=1\right)\\
&= P\left(X_1(w)=1\right)\cdot P\left(X_2(w)=1\right)\\
&= \frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}.
\end{aligned}$$

Otros conceptos construidos a partir de variables aleatorias independientes

Teorema 28 *La esperanza del producto de dos variables aleatorias independientes es igual al producto de sus esperanzas.*

$$E(XY) = E(X)E(Y)$$

Demostración. Sean $X, Y, Z : \Omega \longrightarrow R$ variables aleatorias donde X y Y son independientes y $Z = XY$ entonces $Z(a_i) = X(a_i)Y(a_i)$ para todo $a_i \in \Omega$.

$$\begin{aligned}
E(Z) &= \sum_{i=1}^{n} Z(a_i)\,P(a_i)\\
&= \sum_{i=1}^{n} \left[X(a_i)\,Y(a_i)\right] P(a_i)\\
&= \sum_{i=1}^{n} X(a_i)\,P(a_i) * Y(a_i)P(a_i)\\
&= \sum_{i=1}^{n} X(a_i)\,P(a_i) * \sum_{i=1}^{n} Y(a_i)\,P(a_i)\\
&= E(X)E(Y).
\end{aligned}$$

■

Teorema 29 *La varianza de la suma de dos variables aleatorias independientes es igual a la suma de sus varianzas.* $Var(X+Y) = Var(X) + Var(Y)$

Demostración. Sean $X, Y : \Omega \longrightarrow R$ variables aleatorias independientes

$$\begin{aligned}
Var(X+Y) &= E\left[X+Y-E(X+Y)\right]^2 \\
&= E\left\{[X-E(X)]+[Y-E(Y)]\right\}^2 \\
&= E\left\{[X-E(X)]^2+2\left[X-E(X)\right]\left[Y-E(Y)\right]+[Y-E(Y)]^2\right\} \\
&= E\left[X-E(X)\right]^2+2E\left\{[X-E(X)]\left[Y-E(Y)\right]\right\}+E\left[Y-E(Y)\right]^2 \\
&= Var(X)+Var(Y)+2E\left\{[X-E(X)]\left[Y-E(Y)\right]\right\}
\end{aligned}$$

Resolviendo el producto del término $[X-E(X)]\,[Y-E(Y)]$ se tiene

$$[X-E(X)]\,[Y-E(Y)] = XY - XE(Y) - E(X)Y + E(X)E(Y)$$

obteniendo su esperanza

$$\begin{aligned}
E\left\{[X-E(X)]\,[Y-E(Y)]\right\} &= E\left(XY\right)-2E(X)E(Y)+E(X)E(Y) \\
&= E(XY)-E(X)E(Y).
\end{aligned}$$

Se sabe que X y Y son variables aleatorias independientes entonces $E(XY) = E(X)E(Y)$, por lo que

$$E\left\{[X-E(X)]\,[Y-E(Y)]\right\} = 0$$

Por lo tanto

$$Var(X+Y) = Var(X)+Var(Y)$$

■

Teorema 30 *Sean $X_1, X_2, ..., X_n$ variables aleatorias tal que cada una es independiente de la suma de las anteriores, entonces*

$$Var(X_1+X_2+...+X_n) = Var(X_1)+Var(X_2)+...+Var(X_n).$$

Por inducción establecer que la suma de las primeras n son independientes a n+1.

Corolario 31 *La varianza de la suma de un número finito de variables aleatorias independientes por parejas es igual a la suma de sus varianzas.*

Por la definición de varianza, se tiene:

$$\begin{aligned}
Var\left(\sum_{i=1}^{n} X_i\right) &= E\left[\sum_{i=1}^{n} X_i - E\left(\sum_{i=1}^{n} X_i\right)\right]^2 \\
&= E\left[\sum_{i=1}^{n} \left(X_i - E\left(X_i\right)\right)\right]^2 \\
&= E\left[X_1 + X_2 + ... + X_n - E\left(X_1 + X_2 + ... + X_n\right)\right]^2 \\
&= E\left[\left(X_1 - E\left(X_1\right)\right) + \left(X_2 - E\left(X_2\right)\right) + ... + \left(X_n - E\left(X_n\right)\right)\right]^2 \\
&= E[\left(X_1 - E\left(X_1\right)\right)^2 + \left(X_2 - E\left(X_2\right)\right)^2 + ... + \left(X_n - E\left(X_n\right)\right)^2 \\
&\quad +2\left(X_1 - E\left(X_1\right)\right)\left(X_2 - E\left(X_2\right)\right) + 2\left(X_1 - E\left(X_1\right)\right)\left(X_3 - E\left(X_3\right)\right) \\
&\quad +... + 2\left(X_1 - E\left(X_1\right)\right)\left(X_n - E\left(X_n\right)\right) + 2\left(X_2 - E\left(X_2\right)\right)\left(X_3 - E\left(X_3\right)\right) \\
&\quad +... + 2\left(X_2 - E\left(X_2\right)\right)\left(X_n - E\left(X_n\right)\right) + ... + 2\left(X_{n-1} - E\left(X_{n-1}\right)\right)\left(X_n - E\left(X_n\right)\right)] \\
&= E\left(X_1 - E\left(X_1\right)\right)^2 + E\left(X_2 - E\left(X_2\right)\right)^2 + ... + E\left(X_n - E\left(X_n\right)\right)^2 \\
&\quad +2E\left[\left(X_1 - E\left(X_1\right)\right)\left(X_2 - E\left(X_2\right)\right)\right] + 2E\left[\left(X_1 - E\left(X_1\right)\right)\left(X_3 - E\left(X_3\right)\right)\right] \\
&\quad +... + 2E\left[\left(X_1 - E\left(X_1\right)\right)\left(X_n - E\left(X_n\right)\right)\right] + 2E\left[\left(X_2 - E\left(X_2\right)\right)\left(X_3 - E\left(X_3\right)\right)\right] \\
&\quad +... + 2E\left[\left(X_2 - E\left(X_2\right)\right)\left(X_n - E\left(X_n\right)\right)\right] + ... + 2E\left[\left(X_{n-1} - E\left(X_{n-1}\right)\right)\left(X_n - E\left(X_n\right)\right)\right]
\end{aligned}$$

y por hipótesis son independientes por parejas, es decir todos los productos de la forma

$$2E\left[\left(X_i - E\left(X_i\right)\right)\left(X_j - E\left(X_j\right)\right)\right] = 0$$

para toda $i \neq j$.

Covarianza

Cuando las variables aleatorias no son independientes el término

$$E\left\{\left[X - E(X)\right]\left[Y - E(Y)\right]\right\}$$

no se hace cero, dicho término es conocido como covarianza de X y Y, y se denota como $Cov(X, Y)$. Por lo que

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

Proposición 12 *La varianza de la suma de dos variables aleatorias es igual a la suma de sus varianzas menos dos veces la covarianza.*

$$Var(X + Y) = Var(X) + Var(Y) - 2Cov(X, Y)$$

Observación 32 *Si la $Cov(X, Y) = 0$ no implica que las variables aleatorias son independientes.*

Ejemplo 25 *Sean X y Y variables aleatorias tal que pueden tomar los siguientes valores*

$$X(w) = \begin{cases} 1 \\ 0 \\ -1 \end{cases} \qquad y \qquad Y(w) = \begin{cases} 1 \\ 0 \\ -1 \end{cases}$$

con las probabilidades

$$\begin{aligned} P\left(X(w)=1\right) &= \frac{1}{4},\ P\left(X(w)=0\right)=\frac{1}{2},\ P\left(X(w)=-1\right)=\frac{1}{4} \\ P\left(Y(w)=1\right) &= \frac{1}{4},\ P\left(Y(w)=0\right)=\frac{1}{2},\ P\left(Y(w)=-1\right)=\frac{1}{4} \end{aligned}$$

Conjuntamente (X,Y) solo puede tomar los siguientes valores

$$(1,0),(0,1),(-1,0),(0,-1)$$

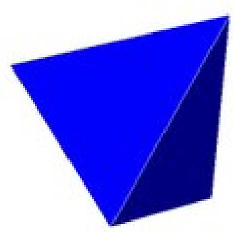
por lo que $E(X)=E(Y)=0$, y como $XY=0$ entonces $E(XY)=0$ lo que implica que $E(XY)=E(X)E(Y)$. Pero X y Y no son independientes pues

$$P(X(w)=0)=P(Y(w)=0)=\frac{1}{2},$$

y

$$P\left[(X(w)=0)\cap\ (Y(w)=0)\right]=0.$$

Ejemplo 26 *Considerese el lanzamiento de dos tetraedros, cada uno numerado en cada lado del uno al cuatro. Sea X la variable aleatoria el número del primer tetraedro y sea Y la variable aleatoria el más grande de los dos números.*



El espacio muestra del experimento es:

$$\Omega = \left\{ \begin{array}{c} (1,1),(1,2),(1,3),(1,4) \\ (2,1),(2,2),(2,3),(2,4) \\ (3,1),(3,2),(3,3),(3,4) \\ (4,1),(4,2),(4,3),(4,4) \end{array} \right\}$$

Las variables aleatorias son funciones de

$X:\Omega\rightarrow R$ donde X puede tomar los valores $X(\omega)=\{1,2,3,4\}$

$Y : \Omega \to R$ donde Y puede tomar los valores $Y(\omega) = \{1, 2, 3, 4\}$

La variable aleatoria X toma el valor 1 cuando sucede cualquiera de los eventos en donde la primera entrada es uno $(1,1)$, $(1,2)$, $(1,3)$, $(1,4)$, por lo que $P(X(\omega) = 1) = \frac{1}{4}$, análogamente para cuando X toma los valores $2, 3$,y 4.

Claramente la variable aleatoria Y depende del valor de la variable aleatoria X por esta razón Y toma el valor 1 cuando sucede unicamente el evento $(1,1)$ y lo hace con probabilidad $\frac{1}{16}$ es decir $P(Y(w) = 1) = \frac{1}{16}$. La variable aleatoria Y toma el valor 2 cuando sucede cualquiera de los siguientes eventos $(2,2), (1,2), (2,1)$ por lo que la $P(Y(w) = 2) = \frac{3}{16}$. Así mismo $P(Y(w) = 3) = \frac{5}{16}$ y $P(Y(w) = 4) = \frac{7}{16}$.

Por lo tanto la esperanza de ambas variables aleatorias es

$$\begin{aligned} E(X) &= 1\left(\frac{1}{4}\right) + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{4}\right) + 4\left(\frac{1}{4}\right) = \frac{5}{2} \\ E(Y) &= 1\left(\frac{1}{16}\right) + 2\left(\frac{3}{16}\right) + 3\left(\frac{5}{16}\right) + 4\left(\frac{7}{16}\right) = \frac{50}{16} \end{aligned}$$

De tal forma que

$$\begin{aligned} E(X^2) &= 1^2\left(\frac{1}{4}\right) + 2^2\left(\frac{1}{4}\right) + 3^2\left(\frac{1}{4}\right) + 4^2\left(\frac{1}{4}\right) = \frac{30}{4}, \\ E(Y^2) &= 1^2\left(\frac{1}{16}\right) + 2^2\left(\frac{3}{16}\right) + 3^2\left(\frac{5}{16}\right) + 4^2\left(\frac{7}{16}\right) = \frac{170}{16}. \end{aligned}$$

Por lo tanto la

$$\begin{aligned} Var(X) &= E(X^2) - E(X) = \frac{30}{4} - \left(\frac{5}{2}\right)^2 = \frac{5}{4} \\ Var(Y) &= E(Y^2) - E(Y) = \frac{170}{16} - \left(\frac{50}{16}\right)^2 = \frac{55}{64} \end{aligned}$$

Y como las variables aleatorias no son independientes entonces

$$Var(X+Y) = Var(X) + Var(Y) - 2Cov(X,Y)$$

para obtener la covarianza es indispensable obtener primero

$$\begin{aligned} E(XY) &= 1*1\left(\frac{1}{16}\right) + 1*2\left(\frac{1}{16}\right) + 1*3\left(\frac{1}{16}\right) + 1*4\left(\frac{1}{16}\right) \\ &\quad +2*2\left(\frac{2}{16}\right) + 2*3\left(\frac{1}{16}\right) + 2*4\left(\frac{1}{16}\right) \\ &\quad +3*3\left(\frac{3}{16}\right) + 3*4\left(\frac{1}{16}\right) + 4*4\left(\frac{4}{16}\right) \\ &= \frac{135}{16} \end{aligned}$$

por lo tanto

$$\begin{aligned} Cov(X,Y) &= E(XY) - E(X)E(Y) \\ &= \frac{135}{16} - \left(\frac{5}{2}\right)\left(\frac{50}{16}\right) = \frac{10}{16} \end{aligned}$$

por lo que

$$\begin{aligned} Var(X+Y) &= Var(X) + Var(Y) - 2Cov(X,Y) \\ &= \frac{5}{4} + \frac{55}{64} - 2\left(\frac{10}{16}\right) = \frac{55}{64}. \end{aligned}$$

Capítulo 5

Teorema central del límite

Como es sabido, fue George Pólya quien dio al célebre resultado el nombre de teorema central del límite para, con ello, enfatizar el papel extremadamente relevante que desempeña en la teoría de la probabilidad. Este teorema explica las razones por las cuales aparecen en todo momento las distribuciones normales, en los diversos campos de aplicación.

El objetivo en esta sección, es enunciar el teorema central del límite, y aunque no se detalla su demostración se hace notar su relación con el teorema integral DeMoivre-Laplace cuya demostración quedó plasmada en el capítulo tres. Ahora serán útiles los conceptos referentes a la esperanza y varianza de una variable aleatoria y veremos cómo este teorema es simplemente una generalización del teorema integral de DeMoivre-Laplace.

Por último, y para hacer énfasis en la importancia de este resultado presentaremos un ejemplo en el cuál se hace notar que es una generalización del teorema integral de DeMoivre Laplace, la diferencia basicamente se reduce a que en el capítulo tres se hicieron los experimentos considerando una moneda, y ahora será con un dado, o con el número de caras que se desee.

El teorema central del límite

El teorema central del límite, uno de los fundamentales en estadística, estudia el comportamiento de la suma de variables aleatorias, cuando crece el número de sumandos, asegurando su convergencia hacia una distribución normal en condiciones muy generales. Este teorema, del cual existen diferentes versiones que se han ido desarrollando a lo largo de la historia, tiene una gran aplicación en inferencia estadística, pues muchos parámetros de diferentes distribuciones de probabilidad, como la media, pueden expresarse en función de una suma de variables. Por otro lado, la suma de variables aleatorias aparece en forma natural en muchas aplicaciones de la economía: salarios, poblaciones, tiempo de espera de servicios, etc.

Como se señaló, el teorema central de límite no es un único teorema, tiene muchas variantes. A continuación estableceremos la versión del teorema que nos interesa abordar.

Teorema 33 (Central del Límite) *Sea $X_1, X_2, ..., X_n$ una sucesión de variables aleatorias independientes identicamente distribuidas, cada una de las cuales tiene esperanza $E(X_k) = \overline{x}$ y varianza finita diferente de cero $Var(X_k) = \sigma^2$. Entonces para todo $\varepsilon > 0$ existe un natural N tal que si $n > N$: para toda $x \in R$*

$$\left| P\left\{ \sqrt{n}\left(\frac{\overline{X}_n - \overline{x}}{\sigma} \right) < x \right\} - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz \right| < \varepsilon.$$

A continuación se usará la notación de los experimentos Bernoulli que definimos en el capítulo tres. Si se considera X como la suma de todas las variables aleatorias

$$X = X_1 + X_2 + ... + X_n = \sum_{k=1}^{n} X_k,$$

que en el caso de los experimentos Bernoulli tendríamos

$$X = \sum_{k=1}^{n} X_k = m,$$

así pues

$$\overline{X}_n = \frac{m}{n}.$$

Por otro lado la esperanza y la varianza de X_k en los experimentos Bernoulli, son las siguientes:

$$E(X_k) = 0 \cdot q + 1 \cdot p = p$$

$$\begin{aligned} Var(X_k) &= E^2(X) - [E(X)]^2 \\ &= 1^2 \cdot p + 0^2 \cdot q - p^2 \\ &= p(1-p) \\ &= pq. \end{aligned}$$

De modo que el teorema anterior puede expresarse de la siguiente forma:

Teorema 34 *Sea $X_1, X_2, ..., X_n$ una sucesión de variables aleatorias independientes con distribución Bernoulli, cada una de las cuales tiene esperanza $E(X_k) = p$ y varianza finita diferente de cero $Var(X_k) = pq = \sigma^2$. Entonces para todo $\varepsilon > 0$ existe un natural N tal que si $n > N$: para toda $x \in R$ Para todo $\varepsilon > 0$ existe un natural N tal que si $n > N$: $x \in R$*

$$\left| P\left\{ \sqrt{n}\left(\frac{\frac{m}{n} - p}{\sigma} \right) < x \right\} - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz \right| < \varepsilon.$$

Como se puede notar, este teorema es una generalización del teorema de DeMoivre-Laplace en su versión integral 18.

Aplicaciones

La idea de esta sección sería mostrar un ejemplo como el que se enunció en las aplicaciones del teorema de DeMoivre-Laplace, la diferencia basicamente se reduce a que anteriormente se hicieron los experimentos considerando una moneda, y ahora lo haremos con un dado, o con el número de caras que se desee.

Se tiene un dado con r caras numeradas como $0, 1, 2, ..., r-1$. Se supone que $p_0,\ p_1, ..., p_{r-1}$ son respectivamente las probabilidades de que en una sola tirada del dado caigan las caras $0, 1, ..., r-1$.

Entonces se tiene que $\sum_{i=0}^{r-1} p_i = 1$.

Se lanza el dado n veces de tal modo que los n lanzamientos son independientes entre ellos.

Para $s = 0, 1, ..., (r-1)n$, se desea saber ¿cuál es la probabilidad $P_n^r(s)$ de que, en los n lanzamientos, la suma de todos los n resultados obtenidos sea exactamente s?

Se tiene ahora que

$$P_n^r(s) = \sum \frac{n!}{k_0!k_1!...k_{r-1}!} p_0^{k_0} p_1^{k_1} ... p_{r-1}^{k_{r-1}}$$

en donde la suma se lleva a cabo sobre todas las $r - adas$ ordenadas $(k_0, k_1...k_{r-1}) \in \overline{n}^r$ tales que $\sum_{i=0}^{r-1} k_i = n$ y $\sum_{i=0}^{r-1} ik_i = s$.

A continuación se presenta un ejemplo de un caso particular en el caso en que

$r = 3$ (número de caras del objeto, en este caso las caras son $0, 1, 2$) ,

$n = 2$ (número de lanzamientos),

$p = \frac{1}{r}$ (con caras igualmente probables).

De lo anterior se puede decir que $s = 0, 1, 2, 3, 4$. Ahora tenemos que buscar las k_0, k_1 y k_2 que cumplan con:

$$\begin{aligned} k_0 + k_1 + k_2 &= 2 \\ 0 \cdot k_0 + 1 \cdot k_1 + 2 \cdot k_2 &= s. \end{aligned}$$

Así para $s = 0$ se tiene que $k_0 = 2,\ k_1 = 0,\ k_2 = 0$, etc. Los resultados se muestran en la siguiente tabla, considerando que para $s = 2$ las dos ecuaciones anteriores tienen dos posibles soluciones.

s	k_0	k_1	k_2
0	2	0	0
1	1	1	0
2	0	2	0
2	1	0	1
3	0	1	1
4	0	0	2

Se calcula ahora $P_n^r(s)$ y se obtiene lo siguiente:

$$\begin{aligned}
P_2^3(0) &= (\frac{1}{3})^2 \\
P_2^3(1) &= 2(\frac{1}{3})^2 \\
P_2^3(2) &= 2(\frac{1}{3})^2 + (\frac{1}{3})^2 \\
P_2^3(3) &= 2(\frac{1}{3})^2 \\
P_2^3(4) &= (\frac{1}{3})^2
\end{aligned}$$

Si las caras no fueran equiprobables se tiene:

$$\begin{aligned}
P_2^3(0) &= p_0^2 \\
P_2^3(1) &= 2p_0p_1 \\
P_2^3(2) &= 2p_0p_2 + p_1^2 \\
P_2^3(3) &= 2p_1p_2 \\
P_2^3(4) &= p_2^2
\end{aligned}$$

donde $p_0 + p_1 + p_2 = 1$.

Programa en Matlab

Para calcular $P_n^r(s)$ se elaboró un programa con ayuda de Matlab, el cual calcula $P_n^r(s)$ dados r, n, y s.

En este programa se supone una figura con tres lados, donde cada una de las caras es equiprobable. El programa pide al usuario que teclee el número de lanzamientos.

```
%(Ahora con una figura de 3 lados)
r=3;        %caras 0,1,2
n=input('numero de lanzamientos');
p=1/r; %que sus caras sean equiprobables


for s=0:((r-1)*n);
    for k0=0:n,
        for k1=0:(n-k0),
            k2=n-k0-k1;
             if (k1+(2*k2))==s,
 P=(factorial(n)/(factorial(k0)*factorial(k1)*factorial(k2)))*(p^k0)*(p^k1)*(p^k2);
                    disp(s)
                    disp(P)
             end
```

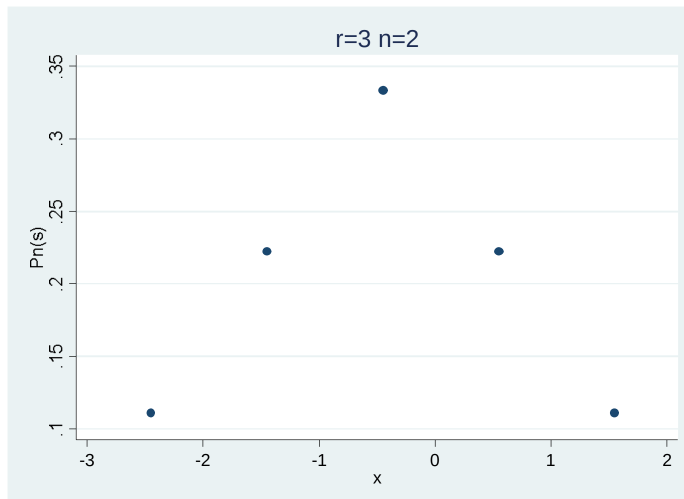
```
            end
        end
end
```

El programa arroja los siguientes resultados con $r = 3$, $n = 2$ y $p = \frac{1}{3}$ que han sido concentradas en la siguiente tabla, y posteriormente se graficaron.

El valor de x fue calculado como lo establece el teorema:

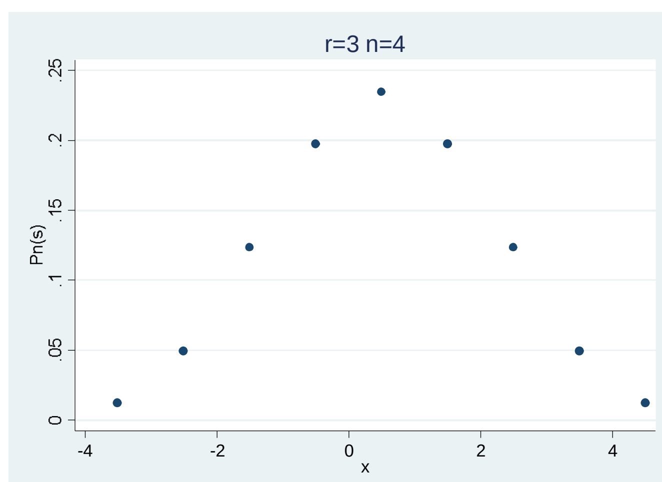
$$x = \frac{s - np}{\sqrt{npq}}.$$

s	x	Pn(s)
0	-2.45	0.1111
1	-1.45	0.2222
2	-0.45	0.3333
3	0.55	0.2222
4	1.55	0.1111



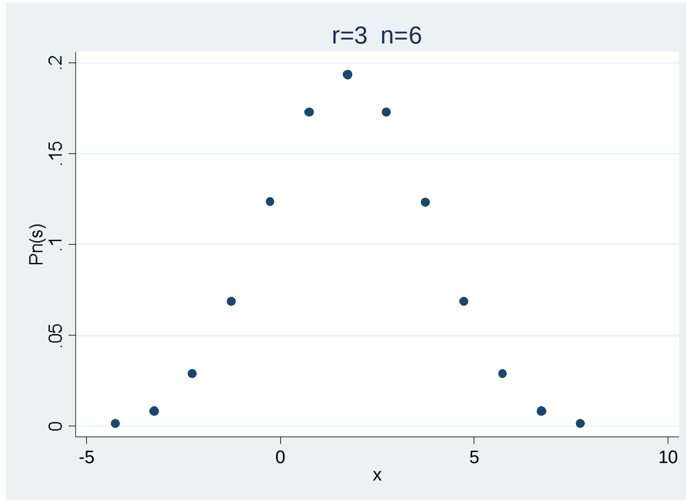
Ahora los cálculos para $r = 3$ y $n = 4$.

s	x	Pn(s)
0	-3.51	0.0123
1	-2.51	0.0494
2	-1.51	0.1235
3	-0.51	0.1975
4	0.49	0.2345
5	1.49	0.1975
6	2.49	0.1235
7	3.49	0.0494
8	4.49	0.0123



Ahora los cálculos para $r = 3$ y $n = 6$.

s	x	Pn(s)
0	-4.26	0.0014
1	-3.26	0.0082
2	-2.26	0.0288
3	-1.26	0.0686
4	-0.26	0.1235
5	0.74	0.1728
6	1.74	0.1935
7	2.74	0.1728
8	3.74	0.1234
9	4.74	0.0686
10	5.74	0.0288
11	6.74	0.0082
12	7.74	0.0014



Lanzamiento de un dado

En este programa se supone que se tiene un dado, donde cada una de las caras es equiprobable. El programa pide al usuario que teclee el número de lanzamientos.

```
%(Ahora con un dado)
r=6;        %caras 0,1,2,3,4,5
n=input('numero de lanzamientos con 6 caras sin suma');
p=1/r; %que sus caras sean equiprobables

for s=0:((r-1)*n);
    for k0=0:n,
        for k1=0:(n-k0),
            for k2=0:(n-k0-k1),
                for k3=0:(n-k0-k1-k2),
                    for k4=0:(n-k0-k1-k2-k3),
                        k5=n-k0-k1-k2-k3-k4;
                         if (k1+(2*k2)+(3*k3)+(4*k4)+(5*k5))==s,
P=(factorial(n)/(factorial(k0)*factorial(k1)*factorial(k2)*factorial(k3)*factorial(k4)
*factorial(k5)))*(p^k0)*(p^k1)*(p^k2)*(p^k3)*(p^k4)*(p^k5);
                                disp(s)
                                disp(P)
                         end
                    end
                end
            end
        end
    end
end
```

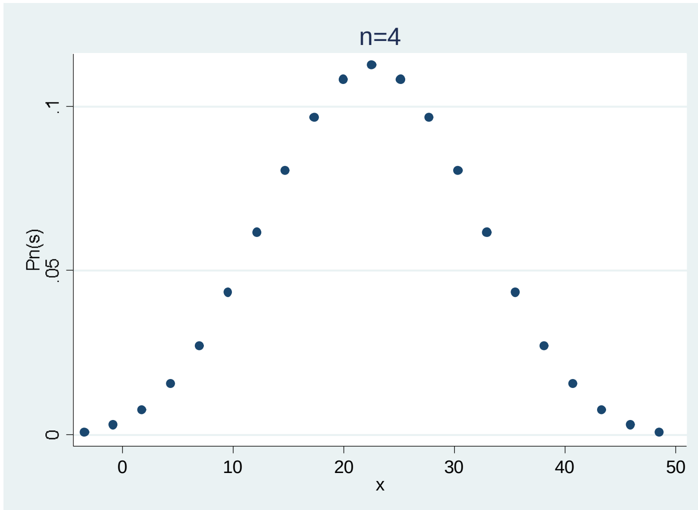
El programa arroja los siguientes resultados con $r = 6$, $n = 4$ y $p = \frac{1}{3}$ porque son equiprobables. que han sido concentradas en la siguiente tabla, y posteriormente se graficaron.

El valor de x fue calculado como lo establece el teorema:

$$x = \frac{s - np}{\sqrt{npq}},$$

los resultados han sido concentrados en la siguiente tabla y posteriormente se graficaron.

s	x	Pn(s)
0	-3.46	0.0008
1	-0.87	0.0031
2	1.73	0.0077
3	4.33	0.0155
4	6.93	0.0271
5	9.53	0.0434
6	12.12	0.0618
7	14.72	0.0805
8	17.32	0.0966
9	19.92	0.1083
10	22.52	0.1127
11	25.11	0.1083
12	27.71	0.0966
13	30.31	0.0805
14	32.91	0.0618
15	35.51	0.0434
16	38.11	0.0271
17	40.70	0.0155
18	43.30	0.0077
19	45.90	0.0031
20	48.50	0.0008



Capítulo 6

Comentarios finales

En el grupo de trabajo de economía matemática y teoria de juejos de la Facultad de Ciencias, y en particular, en el seminario de tesistas, hacemos un esfuerzo por generar material y sugerir ideas didácticas a los estudiantes y profesores, para la enseñanza de temas de matemáticas con aplicaciones a la economía.

Con este espíritu, en un principio este material surge con la idea de estar dirigido a estudiantes de las carreras de economía, sin embargo la primera reflexión que hago con este trabajo, es percatarme de la gran dificultad que existe para diseñar material que lleve de la mano a los estudiantes a abordar temáticas complejas. De tal forma que vale la pena comentar que finalmente más que estar dirigido a estudiantes es una alternativa para profesores.

Considero que las aportaciones sustanciales en este trabajo se encuentran en varios momentos. El primero sería, una alternativa para denotar los espacios finitos de probabilidad, donde se simplifica la notación para no propiciar las dificultades que se generan con los espacios en el sentido de Kolmogorov. Otro momento que me parece importante se encuentra en el significado y demostración del teorema de DeMoivre Laplace, ya que en la demostración se buscó que el concepto de convergencia uniforme se expusiera de tal forma que pueda ser accesible y con la idea de que sea captado del mejor modo posible por el estudiante. Por ejemplo, se encontró una fórmula para calcular la n dada la ε . En cuanto al teorema central del límite, puedo decir que el programa realizado en Matlab es un buen camino para ayudar a entender que este teorema es una generalización del teorema de DeMoivre-Laplace en su versión integral.

Es importante mencionar que yo noto como deficiencia que me hizo falta enriquecer con más ejemplos relacionados a la economía, pero creo también que esto da pie a continuar trabajando con esta propuesta.

Por último quiero hacer énfasis en lo siguiente, después de haber realizando este trabajo, pienso que es importante seguir generando material dedicado a la enseñanza de las matemáticas en las carreras de economía y áreas afines. Ya que considero, que si bien las matemáticas no son la única herramienta para los estudiantes de estas carreras, son de gran utilidad para ayudar a modelar y comprender fenómenos sociales.

Capítulo 7

Bibliografía

[1] Bartle, Robert G. (1995). *The Elements of Integration and Lebesgue Measure.* New York:Wiley -Interscience Publication.

[2] Bartle, Robert G. & Sherbert, Donald R. (2003). *Introducción al análisis matemático de una variable.* Limusa Wiley.

[3] García Álvarez, Miguel Ángel. (2005). *Introducción a la teoría de probabilidad. Primer Curso.* México: Fondo de cultura económica.

[4] García Álvarez, Miguel Ángel. (2005). *Introducción a la teoría de probabilidad. Segundo Curso.* México: Fondo de cultura económica.

[5] Gnedenko, B. V. (1962). *The Theory of Probability.* New York: Chelsea Publishing Company.

[6] Hoel, Paul G. et. al. (1971). *Introduction to Probability Theory.* Boston: Houghton Mifflin Company.

[7] Isaac, Richard. (1995). *The pleasures of probability.* New York: Undergraduate Text in Mathematics. Springer.

[8] Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability.* New York: Chelsea Publishing Company .

[9] Mood, Graybill and Boes. (1974). *Introduction to the Theory of Statistics.* Singapure: Mc Graw Hill.

[10] Papoulis, Athanasios. (1984). *Probability, Random Variables, and Stochastic Processes.* New York: Mc Graw Hill.

[11] Petrov, V. y Mordecki E. (2003).*Teoría de probabilidades.* URSS.

[12] Rincón, Luis. (2007). *Curso intermedio de Probabilidad.* México: Facultad de Ciencias UNAM.

[13] Quesada, V. et al. (2005) *Curso y ejercicios de estadística.* España: Alhambra Longman.

Apéndice A

Conjuntos y Funciones

La matemática moderna está muy relacionada con la teoría de conjuntos, George Cantor matemático alemán inició esta teoría en el siglo XIX, se puede decir que la mayoría de los términos matemáticos están definidos usando el lenguaje de conjuntos, además la esta teoría es fundamental para estudiar la probabilidad. En este apéndice se presentan algunos conceptos básicos relacionados con la notación y las operaciones boleanas.

En cuanto a las funciones, se puede decir que es uno de los conceptos más importantes en matemáticas, y que se usan con mucha frecuencia a lo largo de la tesis. Por lo que en este apéndice, se define de manera general el concepto de función y algunas definiciones relacionadas.

Notación básica de conjuntos

En matemáticas y en la vida diaria siempre encontramos el concepto de conjunto. Se puede hablar del conjunto de caras de un poliedro, el conjunto de estudiantes en un auditorio, el conjunto de puntos en la línea recta, el conjunto de números naturales, etc.

El concepto de conjunto es tan general que es difícil dar una definición que no se reduzca a utilizar alguno de los sinónimos de la palabra conjunto.

Sin embargo podemos dar una definición intuitiva lo suficientemente clara para seguir trabajando en esta dirección.

Entenderemos por "conjunto" una colección o familia de objetos, por ejemplo:

- Los alumnos de la universidad
- Los números primos

Se denotan a los conjuntos con letras mayúsculas $A, B, ...$ y a sus elementos con letras minúsculas $a, b, ...$

A es un conjunto cuando dado un objeto x podamos decir, con toda precisión, si x pertenece o no a A. Y se utiliza la siguiente notación:

$$\begin{aligned} x &\in A \text{ para indicar que “}x\text{ es un elemento de }A\text{”} \\ x &\notin A \text{ para indicar que “}x\text{ no es un elemento de }A\text{”} \end{aligned}$$

Para indicar que A es el conjunto cuyos elementos son $a, b, c, ..., m$ se escribe la lista de elementos entre llaves, separados por comas, es decir

$$A = \{a, b, c, ..., m\}$$

y como con frecuencia no es posible o no es conveniente hacer una lista exahustiva de los elementos de un conjunto; en tal caso, al conjunto de elementos de A lo denotaremos como:

$$\{x|x \in A\}$$

Ejemplos:

1. Sea F el conjunto de alumnos de la Facultad de Economía

   $F = \{x|x \text{ es alumno de la Facultad de Economía}\}$

2. Sea P el conjunto de los números primos

   $P = \{2, 3, 5, 7, ...\}$

3. Sea A el conjunto de los números mayores o iguales que cero y que además satisfacen la ecuación $x^2 - x - 2 = 0$.

   $A = \{x|x \geqq 0, x^2 - x - 2 = 0\}$

4. El conjunto de los números naturales $N := \{1, 2, 3...\}$

5. Las 10 mejores canciones no es un conjunto, porque dada una canción no podríamos decidir si pertenece o no a esta colección. En este caso no existe un criterio único para decidir cuales son las 10 mejores canciones y en la teoría de conjuntos lo que se busca es evitar ambigüedades.

Igualdad

Dos conjuntos son iguales si tienen exactamente los mismos elementos.

Ejemplos:

1. $\{\frac{0}{15}, \frac{4}{4}, 6^0, \sqrt{1}\} = \{0, 1\}$

2. Sea $A = \{x|x^2 - x - 2 = 0\}$ y $B = \{-1, 2\}$. Entonces $A = B$.

3. Sea $A = \{1, 2, 3\}$ y B el conjunto de números enteros mayores que 0 y menores o iguales que 3. Entonces $A = B$.

Inclusión

"A es un subconjunto de B" o "A está contenido en B" si y sólo si todo elemento de A es elemento de B que se denota como $A \subset B$, otra forma de decirlo sería indicar que B contiene a A y se denota $B \supset A$. En este caso $A \neq B$ es decir existe algún elemento en B que no pertenece a A.

Ejemplos:

1. Las vocales forman un subconjunto del abecedario
2. Sea $A = \{-2\}$ y $B = \{x|x^2 - 4 = 0\}$ entonces $A \subset B$ pues -2 es un subconjunto del conjunto solución de la ecuación $x^2 - 4 = 0$.

Observación 35 *Si A incluye todos los elementos de B se puede expresar como $A \subseteq B$ o $B \supseteq A$.*

Conjunto Vacío

El conjunto vacío es el conjunto que no contiene nigún elemento y se denota con el símbolo $\varnothing = \{\ \}$.

Observación 36 *Todo conjunto contiene al conjunto vacío como subconjunto.*

Operaciones con conjuntos

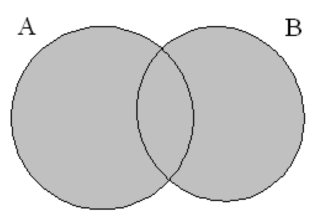
Unión

Definición 15 *Sean A y B dos conjuntos arbitrarios, se define a la **unión**, como*

$$A \cup B = \{x|x \in A \text{ o } x \in B\}$$

consiste en todos los elementos que pertenecen a alguno de los dos conjuntos A o B.

En un diagrama de Venn se puede ver de la siguiente forma



$A \cup B$

Este nuevo conjunto tiene las siguientes propiedades

1. $A \cup \varnothing = A$

2. $A \cup B = B \cup A$

3. $A \cup A = A$

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

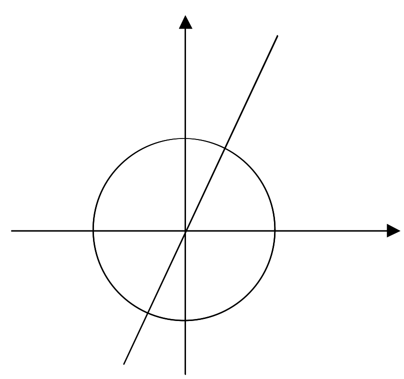
5. $A \subseteq (A \cup B)$ y $B \subseteq (A \cup B)$

6. $A \subset B$ si y solo si $A \cup B = B$

Ejemplos:

1. Sea $A = \{x \in \mathbb{N} \,|x < 10\}$ y sea $B = \{x \in \mathbb{N} \,|x^2 - 144 = 0\}$ entonces
   $A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 12\}$

2. Sea $M = \{(x, y) \in \mathbb{R}^2 \mid y = 3x\}$ y sea $N = \{(x, y) \in \mathbb{R}^2 | \; x^2 + y^2 = 1\}$
   $M \cup N = \{(x, y) \in \mathbb{R}^2 | \; y = 3x \text{ o } x^2 + y^2 = 1\}$
   En este caso, un dibujo describe mejor el nuevo conjunto.



3. Sea $P = \{x \in \mathbb{N} \mid x \text{ es par}\}$ y sea $I = \{x \in \mathbb{N} \mid x \text{ es impar}\}$
   $P \cup I = \mathbb{N}$

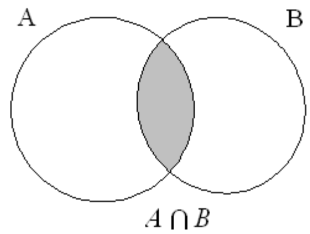
Intersección

Definición 16 *Sean A y B dos conjuntos, se define a la intersección como*

$$A \cap B := \{x | \; x \in A y x \in B\}$$

y consiste de los elementos que pertenecen a ambos conjuntos A y B.

En un diagrama de Venn se puede ver de la siguiente forma



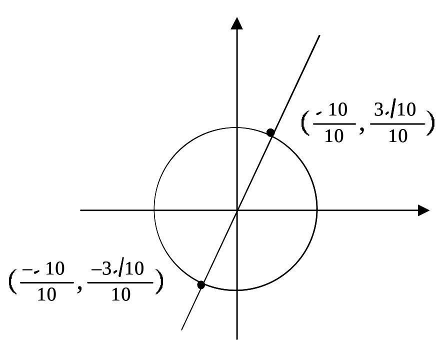
Algunas propiedades de la intersección

1. $A \cap \varnothing = A$
2. $A \cap B = B \cap A$
3. $A \cap A = A$
4. $A \cap (B \cap C) = (A \cap B) \cap C$
5. $(A \cap B) \subset A$ y $(A \cap B) \subset B$
6. $A \subset B$ si y solo si $A \cap B = A$

Definición 17 *Dos conjuntos A y B son ajenos o disjuntos si $A \cap B = \varnothing$*

Ejemplos:

1. Sea $A = \{x \in \mathbb{N} \mid x < 10\}$ y sea $C = \{x \in \mathbb{N} \mid 3 \leq x \leq 12\}$ entonces
   $A \cap C = \{3, 4, 5, 6, 7, 8, 9\}$
2. Sea $M = \{(x, y) \in \mathbb{R}^2 | \ y = 3x\}$ y sea $N = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$
   $M \cap N = \{(x, y) \in \mathbb{R}^2 | \ y = 3x \text{ y } x^2 + y^2 = 1\}$ que serían los puntos de intersección entre la recta y la circunferencia, es decir
   $$M \cap N = \{(\frac{\sqrt{10}}{10}, \frac{3\sqrt{10}}{10}), (\frac{-\sqrt{10}}{10}, \frac{-3\sqrt{10}}{10})\},$$
   graficamente son los puntos siguientes:



3. Sea $P = \{x \in \mathbb{N} \mid x \text{ es par}\}$ y sea $I = \{x \in \mathbb{N} \mid x \text{ es impar}\}$

   $P \cap I = \varnothing$

Dos propiedades importantes que relacionan uniones e intersecciones son las *Leyes Distributivas*

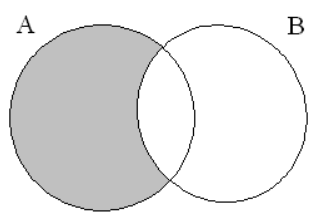
$$\begin{aligned} A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\ A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \end{aligned}$$

Diferencia

Se define una operación más que llamaremos diferencia entre conjuntos y se denota como $A \setminus B$ y son aquellos elementos en A que no estan en B. Es decir

$$A \setminus B = \{x \in A \mid x \notin B\}$$

en diagrama de Venn



$A \setminus B$

Ejemplos:

1. Sea $A = \{x \in \mathbb{N} \mid x < 10\}$ y sea $B = \{x \in \mathbb{Z} \mid x \text{ es par}\}$ entonces

   $A \setminus B = \{1, 3, 5, 7, 9\}$

2. Sea $A = \{x \mid x \text{ es una letra del abecedario}\}$ y sea $B = \{x \mid x \text{ es una vocal}\}$

   $B \setminus A = \varnothing$

Conjunto universal

Todo conjunto puede considerarse como parte de un conjunto mayor en el cual estan considerados todos los elementos que pudieran existir del asunto en cuestión. Para indicar que se están considerando todos los elementos existentes, se emplea el conjunto universal que se representa por $\mathfrak{U}$.

Algunas propiedades inmediatas del conjunto universal son:

1. $A \subseteq \mathfrak{U}$

2. $A \cup \mathfrak{U} = \mathfrak{U}$

3. $A \cap \mathfrak{U} = A$

Ejemplos:

1. El conjunto $\mathbb{N}$ de los números naturales
2. El conjunto de los seres vivientes en la tierra
3. El conjunto de todas las letras del abecedario

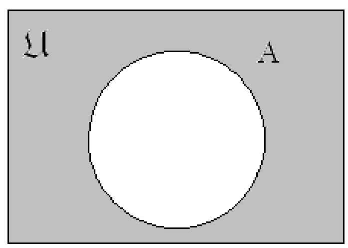
Complemento

Si $A \subset \mathfrak{U}$ podemos definir el conjunto A^c.

A este conjunto se llama el complemento de A (con respecto a $\mathfrak{U}$) y consiste de todos los elementos que estan en $\mathfrak{U}$ pero que no estan en A. Es decir

$$A^c = \mathfrak{U} \setminus A = \{x \in \mathfrak{U}\, |x \notin A\}$$

su diagrama de Venn se puede representar como:



A^c

Este conjunto tiene las siguientes propiedades

1. $(A^c)^c = A$
2. $\varnothing^c = \mathfrak{U}$ y $\mathfrak{U}^c = \varnothing$
3. $A \cap A^c = \varnothing$
4. $A \subset B$ si y solo si $B^{\ c} \subset A^c$
5. Estas propiedades son conocidas como las *Leyes de Morgan*

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

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

Producto Cartesiano

El producto cartesiano de dos conjuntos, se define como

$$A \times B := \{(a,b) \mid a \in A \text{ y } b \in B\}.$$

En general, si $A_1, A_2, ..., A_n$ son conjuntos no vacíos

$$A_1 \times A_2 \times ... \times A_n := \{(a_1, a_2, ..., a_n) \mid a_i \in A_i,\ i = 1, 2, ..., n\},$$

cuando $A_1 = A_2 = ... = A_n = A$ se usa la notación

$$A^n := \{(a_1, a_2, ..., a_n) \mid a_i \in A_i,\ i = 1, 2, ..., n\}.$$

Ejemplos:

1. Sean $A = \{1,2,3\}$, $B = \{a,b\}$. Entonces $A \times B = \{(1,a),(1,b),(2,a),(2,b),(3,a),(3,b)\}$
2. Sea R el conjunto de los números reales. Entonces $R^2 = R \times R = \{(x,y) \mid x \in R,\ y \in R\}$ es el plano real.
3. El conjunto $R^n := \{(x_1, x_2, ..., x_n) \mid x_i \in R,\ i = 1, 2, ..., n\}$ es usado frecuentemente.

Elementos de análisis combinatorio

Cuando se utiliza la definición clásica de probabilidad es necesario contar el número de elementos del evento que interesa y el número de elementos del espacio muestral, lo cual no siempre resulta sencillo, por lo que es necesario utilizar técnicas que faciliten esta tarea.

En este trabajo se revisarán tres técnicas que permiten contar de manera eficiente: el principio fundamental del conteo, las permutaciones y las combinaciones.

Principio fundamental del conteo

Si un evento A puede ocurrir de m formas diferentes y un suceso B puede ocurrir de n formas diferentes, entonces ambos sucesos pueden ocurrir de manera conjunta de mn formas diferentes.

$$\begin{array}{cccc} A & B & & \\ m & n & = & mn \end{array}$$

Ejemplo:

Se va realizar un estudio en el cual se va a clasificar a las personas de acuerdo a su tipo de sangre y a su presión sanguínea. Los tipos de sangre son: O, A, B y AB, y la presión puede ser clasificada como: alta, normal y baja. ¿Cuántas formas se tienen para clasificar a las personas?

Tipo de Sangre	Presión
O	alta
A	normal
B	baja
AB	

Se puede notar que el tipo de sangre O se puede asociar con la presión alta, normal o baja, de manera que hay tres clasificaciones para el tipo de sangre O; de manera análoga se puede ver que también hay tres clasificaciones para A, tres para B y tres para AB. Por lo tanto se tienen 12 posibilidades para clasificar a las personas.

Se tienen cuatro posibilidades para el tipo de sangre y tres para la presión por lo que utilizando el principio fundamental del conteo tenemos que hay $4 * 3 = 12$ posibilidades.

Permutaciones

A cada una de las formas diferentes en que se pueden ordenar los elementos de un conjunto se le llama permutación y generalmente estamos interesados en contar el número de permutaciones diferentes que se pueden hacer con un número determinado de elementos.

Una **permutación** es un arreglo ordenado de elementos.

Ejemplo:

Con los dígitos $0, 1, 2, 3, 4$ y 5 se van a formar números de tres cifras sin que se repitan dígitos en un número. ¿Cuántos números diferentes se pueden formar?

A continuación se muestran algunos de ellos:

012	021	031	041	051
013	023	032	042	052
014	024	034	043	053
015	025	035	045	054

Nótese que hay 20 números que empiezan con 0, por lo tanto hay también 20 números que empiezan con 1 y así sucesivamente, por lo que se tiene que con 6 dígitos se pueden formar 120 números de 3 cifras. La primera cifra puede ser cualquiera de los 6 dígitos, pero sólo quedan 5 para la segunda posición y únicamente quedan 4 posibilidades para ocupar el tercer lugar, por lo que, utilizando el principio funtamental de conteo se tiene: $6 * 5 * 4 = 120$ números diferentes. Así que se tienen 120 permutaciones de tres elementos, lo cual se simboliza de la manera siguiente:

$$P_3^6 = 6 * 5 * 4 = 120$$

donde $P_3^6 =$ es el número de permutaciones de 6 elementos tomando 3 en cada ocasión.

De manera general, cuando se tienen n elementos para formar permutaciones de r elementos, el primer elemento se puede elegir de n formas diferentes, de $n-1$ el segundo, el tercero de $n-2$, y así sucesivamente hasta llegar al lugar r de donde se tienen $n-r+1$ posibilidades, por lo que utilizando el principio fundamental del conte se sigue que:

$$P_r^n = n(n-1)(n-2)(n-3)...(n-r+1)$$

donde:

r es un número menor o igual a n

P_r^n es el número de permutaciones de n elementos tomando r en cada ocasión.

Así, el número de permutaciónes de n elementos tomando r en cada ocasión se puede expresar como:

$$P_r^n = \frac{n!}{(n-r)!}$$

Ejemplo:

En un concurso hay 20 participante, pero sólo hay premios para los cuatro primeros lugares. ¿Cuántas formas se tienen para otorgar los premios?.

Para otorgar el primer lugar hay 20 posibilidades, 19 para el segundo, 18 para el tercero y 17 para el cuarto, por lo que para otorgar los cuatro premios se tienen:

$$P_4^{20} = \frac{20!}{16!} = 20 * 19 * 18 * 17 = 116280$$

posibilidades.

Combinaciones

Hay ocasiones en que resulta de interés seleccionar cierto número de elementos sin que importe el orden, a cada una de las posibles selecciones se le llama combinación.

Una **combinación** es una selección de elementos en la cual no importa el órden.

Ejemplo:

Se van a seleccionar tres elementos del conjunto $U = \{0, 1, 2, 3, 4, 5\}$. Enumerar las diferentes posibilidades que se tienen.

Al seleccionar tres elementos del conjunto U se obtienen las posibilidades siguientes:

$$\begin{array}{ccccc} \{0,1,2\} & \{0,2,3\} & \{0,3,5\} & \{1,2,5\} & \{2,3,4\} \\ \{0,1,3\} & \{0,2,4\} & \{0,4,5\} & \{1,3,4\} & \{2,3,5\} \\ \{0,1,4\} & \{0,2,5\} & \{1,2,3\} & \{1,3,5\} & \{2,4,5\} \\ \{0,1,5\} & \{0,3,4\} & \{1,2,4\} & \{1,4,5\} & \{3,4,5\} \end{array}$$

Cada uno de estos conjuntos representa una combinación de tres elementos tomados del conjunto U. De manera que se tienen 20 combinaciones de 3 elementos tomados de un conjunto de 6 elementos. Esto se puede escribir simbólicamente de la forma siguiente:

$$C_3^6 = 20$$

donde $C_3^6 =$ es el número de permutaciones de 6 elementos tomando 3 en cada ocasión.

Note que con los elementos de la combinación $\{0, 1, 2\}$ se pueden formar las permutaciones que se dan a continuación: $012, 021, 102, 120, 201$ y 210. Pero se tiene que el número de permutaciones de 3 elementos tomando 3 en cada ocasión es:

$$P_3^3 = \frac{3!}{0!} = 3!$$

Así que con cada combinación se pueden formar 3! permutaciones, por lo que si se multiplica el número de combinaciones por 3! se obtiene el número de permutaciones de 6 elementos tomando 3 en cada ocadión, tal como se muestra a continuación:

$$C_3^6 * 3! = P_3^6 = \frac{6!}{3!}$$

y despejando C_3^6, se tiene lo siguiente:

$$C_3^6 = \frac{6!}{3!3!}.$$

El razonamiento anterior se puede hacer de manera general. Si se tienen n elementos de los cuales se van a escoger r (donde r es menor o igual que n), se tienen C_r^n selecciones diferente, pero con cada una de ellas se pueden formar $r!$ permutaciones de r elementos; por lo que si se multiplican estas dos cantidades se obtiene el número de permutaciones de n elementos tomando r en cada ocasión, tal como se muestra a continuación:

$$C_r^n * r! = P_r^n = \frac{n!}{(n-r)!}$$

Luego, el número de combinaciones de n elementos tomando r en cada ocasión es:

$$C_r^n = \frac{n!}{(n-r)!r!}$$

Ejemplo:

Juan tiene doce amigos, pero sólo tiene 4 boletos para invitarlos a un concierto. ¿Cuántas formas diferentes tiene para seleccionar a sus 4 invitados?. Se tienen 12 personas para seleccionar 4, por lo que:

$$C_4^{12} = \frac{12!}{8!4!} = \frac{12 * 11 * 10 * 9}{24} = 495$$

Así que Juan tiene 495 posibilidades para escoger a sus invitados.

Funciones

Definición 18 *Sean A, B dos conjuntos cualesquiera. Una función de A a B es un conjunto f de pares ordenados en $A \times B$, tal que para cada $a \in A$ existe $b \in B$ único, con $(a, b) \in f$ y $(a, b') \in f$, entonces $b = b'$. Al conjunto A se le llama* ***dominio*** *de la funcion f y B es el* ***contradominio****. Al conjunto de los elementos del contradominio que pueden figurar como segunda componente de los pares ordenados de f, se le llama* ***imagen*** *de f.*

Para indicar que f es una función de A a B, se usa la notación

$$f : A \to B.$$

Si $(a, b) \in f$, se acostumbra a escribir

$$b = f(a),$$

también se dice que b es el valor de f en a o que es la imagen de a bajo f.

Algunos tipos de funciones

Definición 19 *Sea $f : A \to B$, si E es un subconjunto de A, entonces la* ***imagen directa*** *de E bajo f es el suconjunto $f(E)$ de B dado por*

$$f(E) := \{f(x) \mid x \in E\}.$$

*Si H es un subconjunto de B, entonces la **imagen inversa** de H bajo f es el subconjunto $f^{-1}(H)$ de A dado por*

$$f^{-1}(H) = \{x \in A \mid f(x) \in H\}.$$

Definición 20 *Sea $f : A \to B$, se dice que f es una función **inyectiva** o **uno a uno** si para todo x_1, x_2 tales que $f(x_1) = f(x_2)$ implica que $x_1 = x_2$.*

De manera equivalente, una función f es inyectiva si y sólo sí $x_1 \neq x_2$ implica que $f(x_1) \neq f(x_2)$.

Definición 21 *Se dice que una función $f : A \to B$ es **suprayectiva** si $f(A) = B$.*

Si una función es suprayectiva también se conoce como **sobreyectiva** o **sobre** simplemente.

Definición 22 *Una función $f : A \to B$ es **biyectiva** si es inyectiva y suprayectiva.*

Función inversa

A partir de una función $f : A \to B$, cuyos pares ordenados son de la forma (a, b), alguien podría preguntarse por otra "función"$g : B \to A$ en la cual se intercambien las entradas de cada par (a, b) por (b, a); en general g **no** es una función, a menos que f sea inyectiva, cuando éste es el caso, g recibe un nombre especial, se conoce como la función inversa.

Definición 23 *Sea $f : A \to B$ una función inyectiva con imagen $R_f \subseteq B$. A la función $g := \{(b, a) \in B \times A \mid (a, b) \in f\}$ con dominio $D_g = R_f$ se le llama **función inversa** de f y se denota por f^{-1}.*

Una función f y su inversa f^{-1} se relacionan de la siguiente forma

$$x = f^{-1}(y) \text{ si y sólo si } y = f(x)$$

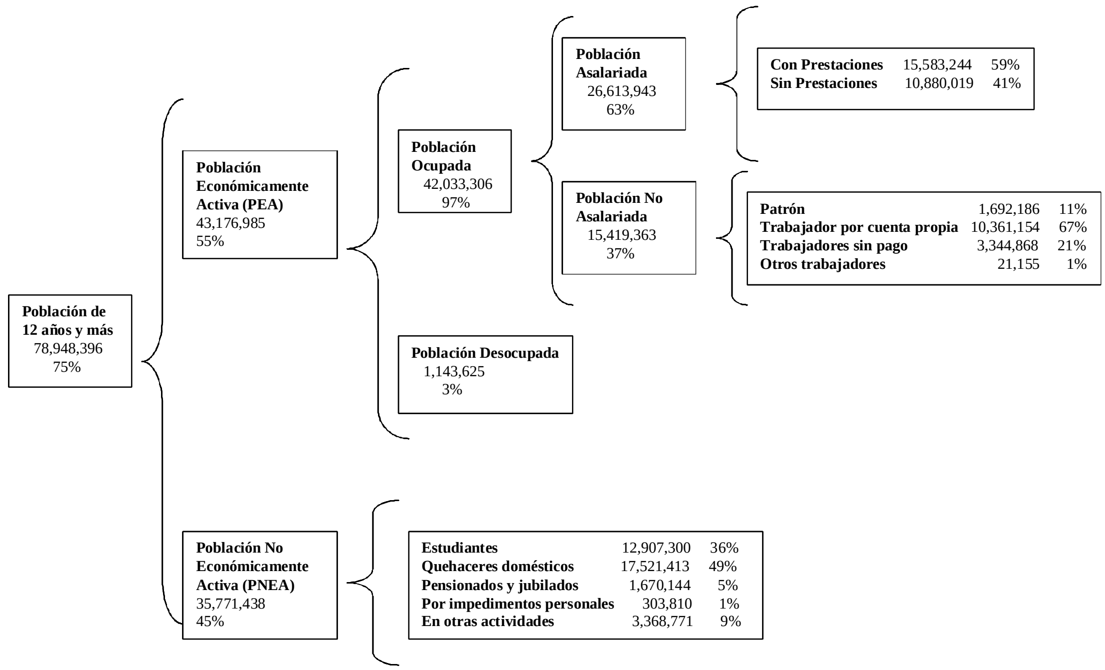
Composición de funciones

Definición 24 *Dadas las funciones $f : A \to B$, $g : B \to C$ se define la **composición** $g \circ f := \{(a, c) \mid \text{para algún } b \in B,\ f(a) = b \text{ y } g(b) = c\}$.*

La definición anterior para la composición de funciones puede ser engorrosa y la manera usual de definirla es

$$(g \circ f)(x) := g(f(x)).$$

Algunas figuras



Fuente: UNITE con cifras del INEGI

Radiografía Ocupacional 2004