UNIVERSIDAD TECNOLÓGICA NACIONAL

Facultad Regional San Nicolás

PROBABILIDAD y ESTADÍSTICA I

UNIDAD Nº2

Licenciatura en Enseñanza de la Matemática

Año 2011

Mg. Lucía C. Sacco

UNIDAD Nº2

Introducción a la probabilidad. Espacios muestrales finitos. Probabilidad condicional.

Modelos matemáticos. Modelos determinísticos versus métodos probabilísticos o aleatorios. Conceptos básicos de probabilidad: espacio muestral, sucesos. Probabilidad empírica o frecuencial.

Probabilidad de ocurrencia de un suceso. Definición clásica de probabilidad y sus limitaciones. Definición axiomática de probabilidad. Consecuencias de los axiomas. Teoremas básicos.

Espacios muestrales finitos. Resultados igualmente probables. Métodos de enumeración: principio de multiplicación, principio de adición, permutaciones, combinaciones, permutaciones cuando no todos los objetos son diferentes.

Probabilidad condicional. Teorema de las probabilidades totales. Teorema de Bayes. Sucesos independientes. Probabilidad condicional e independencia.

Propósitos:

Brindar oportunidades para la construcción de herramientas que permitan:

- Reconocer la importancia de las probabilidades para el ciudadano de hoy.
- La producción del conocimiento aleatorio, analizando diferencias y similitudes con la producción de conocimiento matemático.
- Interpretar cada una de las posibilidades de cálculo de las probabilidades, como así también su correspondiente aplicación.
- Analizar consideraciones generales acerca del aprendizaje de las Probabilidades.
- Interpretar la noción de probabilidad: significados y campos de problemas.
- Analizar secuencias didácticas para el tratamiento de temas de probabilidad.
- Utilizar los principios de la multiplicación y adición en la resolución de probabilidades.
- Resolver problemas relativos a las probabilidades.

Bibliografía sugerida:

- Canavos, George. *Probabilidad y Estadística. Aplicaciones y Métodos*. México. McGraw Hill. 1988.
- Meyer Paul L. *Probabilidad y Aplicaciones Estadísticas*. México. Addison Wesley Iberoamericana .1993.
- Walpole Ronald, Myers Raymond. *Probabilidad y Estadística*. México. Pearson Educación. 1999.
- Apuntes Posgrado: "Estadística Aplicada a la Investigación". Facultad de Ciencias Económicas. Universidad Nacional de Córdoba.

1. Introducción

Generalmente todos tenemos una idea intuitiva de lo que significa el término probabilidad.

Así, cuando decimos es probable que el dólar no siga aumentando, escuchamos en un noticiero que hay un 20% de probabilidad de lluvia, hay poca probabilidad de obtener un producto defectuoso en este proceso de fabricación, es poco probable que en esta central telefónica se reciba más de 100 llamadas entre las 17 y las 17,30 horas, existe poca probabilidad de que este paciente sea diabético, etc., estamos incorporando el término probabilidad en nuestra vida cotidiana como idea de "posibilidad" de ocurrencia de cierto evento.

El concepto de probabilidad ha evolucionado en el transcurso del tiempo. La probabilidad nació en el juego y a los algebristas del siglo XVI, Pacioli, Cardano, Tartaglia, se deben las primeras consideraciones matemáticas profundas a propósito de los juegos de azar.

Los fundamentos del cálculo de probabilidad surgen alrededor del año 1650, cuando sugerido por los juegos de dados, cartas, del lanzamiento de una moneda, se planteó el debate de determinar la probabilidad de ganar la partida.

En 1654 Antoine Gombaud, el caballero de Méré, un jugador compulsivo, pidió a Blaise Pascal que le resolviese el problema del reparto de apuestas cuando se suspendía la partida antes de terminar. Pascal junto con Fermat (ambos matemáticos de esa época) elaboró la solución dando origen al concepto de Probabilidad.

2. Experimentos deterministas y experimentos aleatorios

Consideremos dos listas de experimentos que denominaremos A y B.

Experimentos tipo A

¿Qué ocurrirá si experimentamos poniendo agua en un recipiente y calentándola hasta 100"C? ¿Qué ocurrirá si suelto un vaso que tengo en la mano?

Experimentos tipo B

Si arrojo un dado, ¿qué número aparecerá en la cara superior? Si tiramos una moneda, si seleccionamos al azar una ficha del archivo de participantes del curso a distancia, ¿cuál será la profesión de la persona seleccionada? ¿Cuántas veces debemos arrojarla hasta que salga cara?

Todos aceptarán que, en los experimentos del tipo A, el agua hervirá y el vaso se caerá. En otras palabras, en estos experimentos el resultado es perfectamente previsible.

En cambio, en los experimentos del tipo B, no ocurre lo mismo. Y aunque podríamos llegar a pensar que el resultado es imprevisible, veremos que esta afirmación no es tan taxativa.

Se ignora totalmente el resultado que obtendremos al arrojar un dado, al seleccionar una ficha, o ¿cuántas veces habrá que tirar la moneda?

Es razonable suponer que, al tirar el dado muchas veces, el número tres por ejemplo, saldrá aproximadamente 1/6 de las veces. También podemos pensar, que aproximadamente la mitad de las veces saldrá un número par.

En cuanto a la cantidad de veces que debamos tirar la moneda hasta que salga cara podemos intuir que, probablemente no serán muchas.

En un experimento del tipo B, no podemos predecir de antemano el resultado que obtendremos cada vez que se repita. Pero siempre es posible prever una ley de comportamiento de todos los resultados posibles y, en base a ella, calcular la probabilidad de obtener cada uno de los mismos.

Los experimentos de tipo A, cuyos resultados son previsibles, se conocen con el nombre de **experimentos deterministas**. Los de tipo B, cuyos resultados sólo pueden conocerse estableciendo cierta ley de su comportamiento, se denominan **experimentos aleatorios.**

Estas últimas consideraciones, nos llevan a la siguiente reflexión: e*n los experimentos aleatorios, y sólo en ellos, es posible hablar de probabilidad.*

El problema se presenta cuando debemos cuantificar, medir o calcular estas probabilidades. Este problema, históricamente originado en las preocupaciones de los nobles franceses[1] para tener éxito en los juegos de azar, dio lugar al desarrollo de una teoría matemática **la teoría de probabilidades**.

Como tantas otras teorías matemáticas, la teoría de probabilidades tuvo un desarrollo histórico hasta llegar a la axiomatización. Mediante la axiomatización, se logra sintetizar una definición de probabilidad:

Todo número que cumpla con los tres axiomas de la probabilidad, es una probabilidad.

Si bien esta definición no soluciona el problema de la **forma de medir la probabilidad** para cada experimento concreto, permite en cada caso, medir la probabilidad como resulte más conveniente, siempre que se respeten los axiomas.

Básicamente, existen tres teorías bien conocidas para medir la probabilidad y trataremos de dar una idea somera de ellas por medio de los ejemplos planteados en los experimentos del tipo B.

Cuando arrojamos el dado y queremos calcular, por ejemplo, la probabilidad que salga un tres, utilizaremos la denominada teoría clásica de la probabilidad.

Cuando seleccionamos fichas de inscripción al curso, para medir la probabilidad utilizaremos la teoría frecuencial, donde la probabilidad se mide utilizando la frecuencia relativa de ocurrencia de un evento.

Existirán también situaciones en las que, para cuantificar la probabilidad de cierto desarrollo futuro de los acontecimientos, se debe recurrir a la subjetividad del investigador dando así una medida de probabilidad de tipo subjetivo.

En la **Unidad Nº1**, ya hemos expresado que, en la mayoría de los análisis estadísticos se trabaja con muestras. También dijimos que, el hecho de que un determinado elemento de la población sea tomado en la muestra, debe ser un hecho aleatorio, debido al azar. Una muestra es representativa de la población sólo si podemos asegurar que todos los elementos de la población tuvieron una oportunidad o probabilidad conocida de ser elegidos.

[1] Pascal (1623-1662) y Fermat (1601-1665) consiguieron un auténtico y crucial progreso en la conceptualización de la probabilidad como denota su famosa correspondencia de 1654 donde aparece resuelto el primer problema que el Caballero de Meré le planteó a Pascal y que decía: *Al lanzar dos dados, ¿cuántos lanzamientos son necesarios para tener una probabilidad no menor de 0.5, de conseguir al menos un doble seis?*
Otro problema de interés histórico propuesto a Galileo (1564 – 1642) plantea: *Si se lanzan 3 dados, ¿cuál es la probabilidad de que la suma de los números obtenidos sea, respectivamente 9, 10, 11 ó 12?*
El tratamiento de estos problemas permite apreciar lo importante que resulta explicitar el conjunto formado por todos los posibles resultados de una experiencia aleatoria, como así también un conjunto de posibles resultados. Por tal razón, nos interesa precisar las definiciones de estos conjuntos y llegar a la definición formal de probabilidad.

Como los resultados de las muestras son valores aleatorios, generalizar las conclusiones de la muestra a la población, produce una situación de incertidumbre.

Ahora bien, si logramos establecer una cierta ley de comportamiento de los resultados muestrales, la confianza de la generalización puede ser medida en términos de probabilidad.

En esta **Unidad Nº2** se hace énfasis, a la manera de establecer leyes de comportamiento de los fenómenos aleatorios. A partir de ellas, podremos medir la incertidumbre que produce generalizar las conclusiones de la muestra a una cierta población; principio fundamental de la inferencia estadística.

Podemos entonces anticipar que la probabilidad es el instrumento adecuado para medir el "error" que se comete al inferir resultados de una muestra, a la población cuando los elementos que la componen han sido seleccionados aleatoriamente y con una cierta probabilidad conocida.

3. Definiciones de Probabilidad

3.1 Espacio muestral y evento

Antes de definir probabilidad, introduciremos algunos conceptos previos.

Comenzamos apelando a los juegos de salón, en los que el azar desempeña un papel esencial.

Dado el siguiente ejemplo:

Ejemplo 1

En muchos juegos de naipes se utiliza una baraja de 52 cartas; dividida en dos colores: rojo y negro, y cuatro palos: corazón, diamante, trébol y pique. Dentro de cada palo hay 13 cartas: 1, 2, 3, ... 10, J, Q, K.

El mazo de naipes, antes de comenzar el juego, se mezcla varias veces a los fines de colocar las cartas en un orden aleatorio. Esto significa que todos los posibles ordenamientos de las cartas son igualmente probables. Con este procedimiento, al distribuir las 52 cartas del mazo, todas tienen la misma probabilidad de ser, por ejemplo, la primera carta extraída.

Es fácil deducir que para el jugador tienen mucha importancia la expresión aleatoria e igualmente probable, pues son garantía de que se juega sin trampas.

La mezcla de las cartas del mazo puede ser considerada como una manera física de aproximarse a la aleatoriedad.

En sus comienzos, el cálculo de probabilidades estaba solamente relacionado con los juegos de azar, y básicamente sustentado en que todos los resultados de un juego, tenían la misma probabilidad de ocurrir.

Esta situación dio origen a una importante obra, escrita por el matemático Laplace quien desarrolló una teoría de probabilidades basada en que los resultados de un cierto experimento tenían una equiprobabilidad de ocurrir.

Además de otras ventajas, esta teoría tiene el gran mérito de haber sentado las bases para la elaboración de otras más recientes.

Antes de abordar estas teorías, daremos algunas definiciones necesarias para facilitar su comprensión.

El cálculo de probabilidades es la teoría matemática que construye modelos para la descripción y análisis de los eventos aleatorios.

¿Qué entendemos por evento aleatorio?

Un evento aleatorio, es un resultado de un experimento aleatorio

Recordando los conceptos iniciales sobre experimentos aleatorios, podemos enunciar:

Un experimento aleatorio es una operación realizada un cierto número de veces, bajo las mismas condiciones de experimentación.

Un resultado no puede preverse cuando se realiza una sola vez pero, si se repite un número grande de veces, los resultados del experimento responderán a una ley de comportamiento regular y previsible.

Ejemplos 2

Tirar una moneda al aire, es un experimento aleatorio. La obtención de una cara es un evento aleatorio pues antes de que la moneda toque el suelo no se puede predecir si el resultado será cara o cruz.

Extraer una carta de un mazo de naipes es un experimento aleatorio, y la obtención de, por ejemplo un trébol, es un evento aleatorio.

Los resultados posibles de estas operaciones, se consideran resultados igualmente probables. Esta idea intuitiva parte de suponer que el mazo tiene un orden aleatorio, que la moneda es perfecta y que el dado no está cargado.

Se puede estar interesado en conocer la probabilidad de que la carta que se extrae tenga alguna propiedad específica: que sea un trébol.

El evento que la primera carta extraída sea trébol corresponde a un conjunto de 13 resultados posibles: as, dos, tres, ..., K dé trébol.

Pero entonces, ¿cuál será la probabilidad de extraer un trébol? Dado que hay 13 tréboles y todas las cartas son igualmente probables, parece razonable considerar la probabilidad de extraer un trébol como 13/52 = 1/4.

Esto nos lleva a enunciar **la definición clásica de probabilidad**.

3.2 Definición clásica de Probabilidad

Si A es un suceso del espacio muestral asociado a una experiencia aleatoria, no podemos decir a priori si A ocurrirá o no, al realizar la experiencia.

Por tal razón interesa asociar con cada suceso del espacio muestral un número que mida, de alguna manera, la posibilidad que tiene A de ocurrir. Esta tarea nos conduce al concepto de probabilidad.

Si consideramos el suceso A: sale un número par al tirar un dado, resulta razonable considerar que ese suceso tiene una posibilidad de ocurrir igual a 0,5 en virtud de que el 50 % de los resultados posibles pertenecen al suceso A.

Este razonamiento es factible en virtud de que:

- ♦ podemos calcular el porcentaje de resultados que hay en A, con respecto al espacio muestral S, ya que éste tiene un número finito de elementos.
- ♦ suponemos que cada resultado de S tiene la misma posibilidad de ocurrir.

Esta última consideración es admisible si el dado está bien construido (dado equilibrado) y en consecuencia las seis caras tienen la misma posibilidad de salir.

El primer intento de definir con rigor matemático el concepto de probabilidad es debido a Laplace (1812) quien dio la definición que se conoce con el nombre de definición clásica, y dice:

La probabilidad de un suceso es igual al cociente entre el número de resultados favorables y el número total de resultados posibles, siempre y cuando todos los resultados tengan igual posibilidad de ocurrir.

Esta definición, incluso para la misma época, resulta circular y restrictiva, y sólo ofrece un método práctico de cálculo para algunas situaciones.

Ejemplo 3

Un bolillero contiene 20 bolillas numeradas de 1 a 20. Si se elige al azar una bolilla, ¿cuál es la probabilidad de que el número de la bolilla elegida sea divisible por 3?

Si notamos con A al suceso: "el número de la bolilla elegida al azar es divisible por 3" resulta A = {3, 6, 9, 12, 15, 18}.

Siendo el conjunto de todos los resultados posibles S = {1, 2, 3,........20}, la probabilidad del suceso A es igual a $\frac{6}{20}$, y notamos P(A) = $\frac{6}{20}$

Para calcular esta probabilidad tuvimos que suponer que las apariciones de cada bolilla tienen la misma probabilidad de ocurrir (ley de comportamiento).

Es importante destacar, en este punto, que la teoría clásica no se aplica únicamente a los juegos de azar sino que toda la teoría del muestreo de poblaciones finitas, está basada en la versión Laplaciana de la probabilidad, versión actualmente modificada o enriquecida por el enfoque Neo Bayesiano.

3.3 Definición de probabilidad empírica o frecuencial

Hay situaciones en las cuales no se puede aplicar la noción de un número definido de casos igualmente probables.

Ejemplo 4

Si una moneda no es perfecta, la probabilidad de obtener cara al lanzarla, no es 1/2, debido a que los dos resultados posibles no son igualmente probables.

En estos casos, se puede aplicar la teoría de probabilidad debida a von Mises, conocida con el nombre de **teoría frecuencial de la probabilidad**.

Esta teoría expresa que, si se repite un experimento aleatorio un número bastante grande de veces, la probabilidad de un evento en particular puede asimilarse a la frecuencia relativa.

En otras palabras, la ley de comportamiento de los fenómenos aleatorios surge de la distribución de frecuencias con que se presentan.

Este enfoque del cálculo de la probabilidad, a partir del concepto de frecuencia relativa, es el más ligado a los hechos de la vida cotidiana y a la aplicación práctica de la estadística.

Supongamos que realizamos el experimento A un número n muy grande de veces. Cada resultado individual del experimento será impredecible puesto que la obtención del resultado depende estrictamente del azar. Al repetir el experimento una elevada cantidad de veces, la frecuencia relativa se estabiliza entonces a este valor se lo denomina probabilidad de A:

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

Es lógico pensar que, si tiramos la moneda un número n de veces, donde $n = 100$, a pesar del comportamiento irregular de los resultados individuales, los resultados promedios o globales mostrarán una sorprendente regularidad.

Al finalizar el experimento, podremos cuantificar la probabilidad del evento que consiste en la caída de la moneda mostrando la faz "cara". Ese valor estará dado por el cociente entre, la cantidad de veces que la moneda cayó en esa posición y la cantidad n de veces que se la arrojó.

Como ya hemos visto en la **Unidad Nº1**, el número de veces que se presenta un evento es la frecuencia absoluta f_a y la frecuencia relativa $f_r = f_a/n$ representa la proporción de veces que se presenta un evento en particular, en las n repeticiones del experimento.

La experiencia indica que la frecuencia relativa tiende a estabilizarse para grandes valores de n.

Es decir, podemos establecer una ley de comportamiento de los fenómenos aleatorios a través de muchas repeticiones del experimento.

Entonces, si definimos al evento A = salida de cara, la frecuencia relativa $f_r(A)$ varía ampliamente para valores pequeños de n. Pero, al aumentar el número n de repeticiones del experimento, la proporción de "caras" se va estabilizando alrededor de un cierto valor límite o ideal muy próximo a 1/2.

Al repetir muchas veces un mismo experimento en condiciones uniformes, la estabilidad de las frecuencias relativas se presenta de manera casi permanente.

Esta situación, fuertemente ligada a la experiencia, llevó a von Mises a enunciar la teoría frecuencial de la probabilidad razonando de la siguiente manera:

Podemos asociar un número P(A) como probabilidad de cada evento A surgido por medio de la realización de un experimento aleatorio E, de manera tal que $f_r(A)$, será aproximadamente igual a P(A) si se considera una larga serie de repeticiones del experimento.

Diremos que P(A) es la probabilidad del evento A en el experimento E.

No debemos olvidar aquí, que la frecuencia relativa y la probabilidad, al vincularse, establecen una relación entre una situación de experimentación real (frecuencia realmente observada) y un modelo conceptual o ideal (frecuencia teórica o esperada).

Por eso, esta asimilación de la probabilidad a las frecuencias relativas sólo es posible ante experimentos que se puede repetir indefinidamente. Y aún así, siempre cabrá la duda respecto de la magnitud que debe tener n para que esta aproximación se produzca.

Ejemplo 5

Podemos suponer que cierta empresa tiene un archivo con todos los reclamos de garantía efectuados por los clientes durante un año. Se decide tomar una muestra de n = 70 reclamos y clasificarlos por categoría de defectos. (Partimos de suponer que cada reclamo de garantía contempla una y sólo una de las categorías de defectos).

La extracción de los reclamos puede ser considerada un experimento aleatorio pues el resultado individual registrado en cada ficha no puede predecirse hasta que la misma sea seleccionada y leída. Este experimento aleatorio se repite un número n = 70 veces bajo condiciones uniformes. Los resultados obtenidos se resumen en la siguiente tabla:

Categoría de defecto	Cantidad de reclamos n_i	Frecuencia relativa fr_i
Menor	38	0,543
Apenas serio	20	0,286
Serio	8	0,114
Mayor	4	0,057
Total	**70**	**1,000**

La frecuencia relativa es la cantidad de veces que ocurre una categoría de reclamo en especial dividida por el total de reclamos:

$$f_{r_i} = {f_{a_i}}/{n}$$

Es lícito entonces, asimilar esta frecuencia relativa con la probabilidad de que ocurra un evento en particular.

En este experimento, un evento podría ser, por ejemplo, la ocurrencia de un defecto mayor.

Definimos, entonces, al evento:

M = se registra un reclamo por un defecto mayor

En base a la experiencia realizada podemos decir, entonces, que existe una probabilidad de un 6% (0,057) de que la empresa reciba reclamos de garantía por defectos mayores.

En símbolos,

$$P(M) = \frac{4}{70} = 0.057 \longrightarrow \text{ aproximadamente } 0{,}06$$

Este dato puede ser importante para el Departamento de Costos ya que los reclamos de garantía por defectos mayores son los que más inciden en los presupuestos y, como puede observarse, tienen una baja probabilidad de ocurrir.

3. 5 Definición de probabilidad subjetiva

Se refiere a la probabilidad de ocurrencia de un suceso basado en la experiencia previa, la opinión personal o la intuición del individuo. En este caso después de estudiar la información disponible, se asigna un valor de probabilidad a los sucesos basado en el grado de creencia de que el suceso pueda ocurrir.

4. Axiomas de probabilidad

En 1993, el enfoque axiomático de la probabilidad fue formalizado por el matemático ruso A. N. Kolmogorov. La base de este enfoque está formalizada en tres axiomas, en las que la probabilidad de un evento A en el experimento aleatorio E, es el valor numérico P(A) que satisface los mismos:

Primer axioma

1) Si A es un evento, luego P(A) ≥ 0 para todo A

Segundo axioma

2) Si S representa al conjunto de todos los resultados posibles de un experimento aleatorio (espacio muestral), luego P(S) = 1

Tercer axioma

3) P (A1 U A2 U...) = P (A1) + P (A2) + ... si A1, A2, ... es una sucesión finita o infinita de eventos incompatibles o excluyentes.

Ejemplo 6

Supongamos que un experimento aleatorio consiste en tirar un dado perfecto, sólo una vez.

El conjunto formado por todos los resultados posibles de este experimento, S, será:

W = {1, 2, 3, 4, 5, 6}

Aplicando la definición clásica de probabilidad, tenemos:

P (de que salga el 1) = P (1) = 1/6

P (de que salga el 2) = P (2) = 1/6

P (de que salga el 3) = P (3) = 1/6

P (de que salga el 4) = P (4) = 1/6

P (de que salga el 5) = P (5) = 1/6

P (de que salga el 6) = P (6) = 1/6

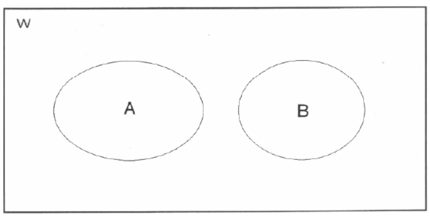
Como podemos observar, todas las probabilidades calculadas son números mayores o iguales a 0.

Veremos ahora como podemos calcular la probabilidad de S.

Como estamos tirando una sola vez un dado, si sale el 1 no puede salir al mismo tiempo el 2 o el 3 o el 4, etc. Decimos entonces, que se trata de eventos mutuamente **excluyentes**.

Eventos mutuamente excluyentes son aquellos que no pueden presentarse conjuntamente.

Apelando a la teoría de conjuntos y utilizando el conocido diagrama de Venn, la representación gráfica de dos eventos A y B mutuamente excluyentes, será:



Así, A U B significa que, al efectuar un experimento, aparece el **evento A** o el **evento B** o ambos, donde el término "o" indica la operación matemática de la suma. Cuando A y B son mutuamente excluyentes no pueden aparecer juntos, por lo cual la unión implica la aparición de A o de B pero no la de ambos a la vez.

Generalizando:

Al definir el conjunto total de resultados posibles hemos utilizado el concepto de aparición de un 1 o de un 2, etc. y, como todos estos eventos son mutuamente excluyentes, podemos escribir:

P (S) = P (1 ó 2 ó 3 ó 4 ó 5 ó 6) =

= P (1 U 2 U 3 U 4 U 5 U 6) =

= P (1) + P (2) + P (3) + P (4) + P (5) + P (6) =

= 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 =

= 6/6 = 1

Al calcular P (S) hemos verificado el segundo axioma P (S) = 1

4.1 Consecuencias de los axiomas

Antes vamos a presentar y a explicar, cuatro consecuencias importantes que se desprenden de los tres axiomas enunciados:

1) P (φ)= 0 (φ es un evento imposible)

2) P (A) = 1 – P ($\overline{A}$) ($\overline{A}$ es el evento complemento de A)

3) 0 ≤ P(A) ≤ 1 cualquiera sea el evento A

4) P (A1 U A2) = P (A1) + P (A2) – P (A1 ∩ A2) cualesquiera sean A1 y A2

Primera Consecuencia

Así como existe un conjunto que comprende a todos los resultados posibles de un experimento aleatorio (S) conocido matemáticamente como conjunto universal, existe también un conjunto que no tiene ningún elemento o resultado posible. Es el conjunto vacío y se lo simboliza con φ.

Ejemplo 7

La posibilidad de obtener un 7 o cualquier otro valor mayor a 6 al tirar un dado, es un resultado imposible de obtener.

Un evento imposible es aquel que no tiene ningún resultado favorable dentro del conjunto de resultados posibles de un experimento aleatorio.

Al utilizar la definición clásica de probabilidad, se establece que la probabilidad de φ es 0. En símbolos

P (φ) = 0

Segunda Consecuencia

Veremos a continuación el concepto de evento complemento.

El evento complemento $\overline{A}$ de un evento A es el evento que consiste de todos los resultados que no contiene el evento A.

Por definición, los eventos $\overline{A}$ y A son mutuamente excluyentes.

Entonces $P(\overline{A} \text{ o } A) = P(\overline{A} \cup A) = P(\overline{A}) + P(A) = 1$

Luego, despejando $P(\overline{A})$, se tiene $P(\overline{A}) = 1 - P(A)$

Tercera Consecuencia

En cuanto a la consecuencia 3), podemos decir que la probabilidad es un número que varía entre 0 y 1 ya que, considerando las dos situaciones extremas, tenemos:

$$P(\varphi) = 0 \text{ y } P(S) = 1$$

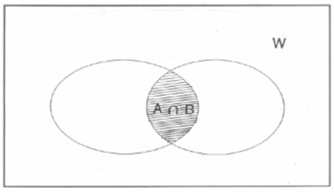
Siendo φ el evento imposible y W el evento seguro o cierto.

Cuarta Consecuencia

Veremos a continuación qué ocurre cuando debemos calcular la probabilidad de eventos que no son mutuamente excluyentes.

Eventos no mutuamente excluyentes son aquellos que tienen resultados en común.

Utilizando el diagrama de Venn, la representación gráfica de dos eventos A y B que no son mutuamente excluyentes, es la siguiente:



Si observamos el diagrama de Venn, la parte sombreada representa la cantidad de resultados de un experimento en el que se presentan conjuntamente los eventos A y B.

La presentación conjunta de estos eventos se simboliza como $A \cap B$, siendo $\cap$ el símbolo de intersección de la teoría de conjuntos.

Trataremos ahora de encontrar la probabilidad de A o B cuando los eventos A y B no son mutuamente excluyentes.

Al hablar del concepto de unión de eventos definimos que $A \cup B$ representa la aparición de los eventos A o B, o ambos a la vez, al efectuar un experimento.

En este caso $P(A \text{ o } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$

El primer término de la suma representa la probabilidad de ocurrencia del evento A y el segundo término representa la probabilidad de ocurrencia del evento B.

Pero cada uno, incluye los resultados de aparición conjunta de ambos eventos (A y B). Por este motivo, el tercer término resta una vez la probabilidad de la intersección que, de otro modo, sería sumada dos veces.

Ejemplo 8

Supongamos que en una muestra de 30 niños se obtuvieron los siguientes resultados referidos a peso y altura:

Calcular la probabilidad de que un niño de esta muestra sea alto o flaco.

	Alto (B_1)	**Bajo (B_2)**	**Total**
Gordo (A_1)	12	6	18
Flaco (A_2)	8	4	12
Total	**20**	**10**	**30**

P (alto o flaco) = P (de que sea alto o de que sea flaco o de que sea alto y flaco)

P (alto o flaco) = $P(B1 \cup A2) = P(B1) + P(A2) - P(B1 \cap A2)$

$$= \frac{20}{30} + \frac{12}{30} - \frac{8}{30} = \frac{24}{30} = 0{,}80$$

Luego, la probabilidad de que un niño elegido al azar sea alto o flaco es igual a 0,80. A esta probabilidad se la denomina **probabilidad total.**

Es posible, realizar una correspondencia de términos entre la teoría de conjuntos y la teoría de probabilidad:

En teoría de probabilidad	En teoría de conjuntos
Evento	Conjunto
Resultado	Elemento
Mutuamente excluyentes	Disjuntos
A o B	A U B (A unión B)
A y B	A $\cap$ B (A intersección B)
A	A, $\overline{A}$ (complemento de A)
Evento nulo (ϕ)	Conjunto vacío

5. Probabilidad Condicional

Existen otras situaciones que se pueden presentar al calcular la probabilidad de la intersección; es decir, la probabilidad de que A y B se presenten conjuntamente.

Seguimos con el ejemplo anterior, la muestra de 30 niños que se clasifican en peso y altura.

Primera Situación

Una primera situación es la que se plantea cuando simplemente debemos calcular la **probabilidad de la aparición conjunta de dos eventos.**

- Calcular la probabilidad de que un niño sea a la vez gordo y alto. Observemos en la tabla elaborada al analizar este caso, que la celda donde se reúnen los resultados posibles de ser gordo y alto es la correspondiente al cruce de los eventos A1 y B1 cuya frecuencia es 12.

P (gordo y alto)= P (A1 ∩ B1)

$$P(A1 \cap B1) = \frac{12}{30} = 0{,}40$$

O queremos calcular las **probabilidades marginales** o del **margen de la tabla**:

- Calcular la probabilidad de que un niño sea gordo:

$$P(A1) = \frac{18}{30} = 0{,}60$$

- Calcular la probabilidad de que un niño sea flaco será:

$$P(A2) = \frac{12}{30} = 0{,}40$$

Segunda Situación

Considerando únicamente la variable peso, es posible dividir la muestra de 30 niños en dos submuestras: niños gordos y niños flacos.

En la tabla de frecuencias, tendremos:

	Alto (B_1)	Bajo (B_2)	Total	
Gordo (A_1)	12	6	18	(Submuestra 1)
Flaco (A_2)	8	4	12	(Submuestra 2)
Total	**20**	**10**	**30**	

- Considerando la submuestra de niños flacos, determinar la probabilidad de que, además, sean altos.

Los resultados favorables a este evento en particular, serán los que consideran, simultáneamente, los dos eventos: niños flacos y niños altos.

Es decir que, los casos posibles ya no serán los 30 niños de la muestra sino los 12 niños correspondientes a la submuestra determinada por la propiedad de ser flacos.

Al tener una información previa (ya sabemos que el niño es flaco), la muestra total de 30 niños se restringe a la submuestra que posee la propiedad que actúa como condicionante.

La probabilidad de que un niño sea alto, está condicionada al hecho de saber a priori, que es flaco.

En símbolos $P(B1/A2)$ y esto se lee: probabilidad de que ocurra B1 dado que ya ocurrió A2. Considerando que una probabilidad se calcula como el cociente entre resultados favorables y resultados posibles, tendremos $P(B1/A2)=\frac{8}{12}$

Si dividimos numerador y denominador del segundo miembro por él total de elementos de la muestra, tendremos $P(B1/A2)=\frac{8}{12}=\frac{\frac{8}{30}}{\frac{12}{30}}$

El valor 8/30 es la probabilidad de que se den conjuntamente los eventos B1 y A2 —› P (B1 ∩ A2) y el valor 12/30 es la probabilidad del evento A2—› P (A2)

Reemplazando en la fórmula anterior, tenemos $P(B1/A2)=\frac{P(B1\cap A2)}{P(A2)}$

Esta probabilidad se conoce con el nombre de **probabilidad condicional**. Se denomina condicional porque el evento A2, el que ya ocurrió, actúa como condicionante modificando la probabilidad de ocurrencia de B1.

Tercera Situación

Razonando de la misma manera, podríamos pensar:

- Calcular la probabilidad de que un niño sea gordo condicionado al hecho de saber que es alto.

La tabla de frecuencias, quedaría dividida en las siguientes submuestras:

	Alto (B_1)	Bajo (B_2)	Total
Gordo (A_1)	12	6	18
Flaco (A_2)	8	4	12
Total	**20**	**10**	**30**
	(Submuestra 1)	(Submuestra 2)	

La probabilidad condicional buscada es $P(A1/B1)$

Entonces, los casos posibles se limitan a aquellos niños que son altos (20 niños). Los casos favorables, comprenderán a los niños que se ubicaron en la categoría altos y gordos (12 niños). Luego, la probabilidad buscada es:

P (de que un niño sea gordo sabiendo de antemano que es alto) =

$$P(A1/B1) = \frac{P(A1 \cap B1)}{P(B1)} = \frac{\frac{12}{30}}{\frac{20}{30}} = 0{,}60$$

En síntesis:

$$\boxed{P(A/B) = \frac{P(A \cap B)}{P(B)}}$$

De la igualdad anterior, por un pasaje matemático de términos, tenemos:

$$\boxed{P(A \cap B) = P(A/B)\,.\,P(B) = P(B/A)\,.\,P(A)}$$

Esta expresión se conoce como *el teorema de multiplicación* de probabilidades.

Es posible aplicar este teorema al cálculo de las probabilidades de la ocurrencia simultánea de dos sucesos A y B.

Para el **Ejemplo 8**, veremos que se obtiene el mismo resultado que cuando calculamos directamente P (A1 ∩ B1):

$$\boxed{\begin{gathered} P(A1 \cap B1) = P(A1/B1)\,.\,P(B1) = P(B1/A1)\,.\,P(A1) \\ = \frac{12}{20}\,.\,\frac{20}{30} = \frac{12}{18}\,.\,\frac{18}{30} = \frac{12}{30} = 0{,}40 \end{gathered}}$$

5.1 Eventos independientes

Otro concepto muy importante es el relacionado con la independencia de eventos.

Ejemplo 9

Treinta y cinco hombres de una muestra, se clasificaron de acuerdo a sus hábitos de fumar y considerando si tienen o no frecuentes dolores de cabeza.

Los resultados obtenidos fueron:

	Fuman (A_1)	**No fuman** (A_2)	**Total**
Con frecuentes dolores (B_1)	12	8	20
Sin frecuentes dolores (B_2)	9	6	15
Total	**21**	**14**	**35**

En base a esta tabla de frecuencias:

- la probabilidad condicional de seleccionar una persona que tenga frecuentes dolores de cabeza habiendo fumado será:

$$P(B1/A1) = \frac{P(A1 \cap B1)}{P(A1)} = \frac{12}{21} = 0{,}57$$

- la probabilidad de seleccionar una persona que tenga dolores de cabeza y no haya fumado es:

$$P(B1/A2) = \frac{P(A2 \cap B1)}{P(A2)} = \frac{8}{14} = 0{,}57$$

De lo calculado precedentemente se puede deducir que la probabilidad de que un individuo tenga frecuentes dolores de cabeza no depende del hábito de fumar.

Esta probabilidad es la misma ya sea que la persona fume o no. En consecuencia, que una persona padezca dolores de cabeza es independiente del hecho de fumar.

Podemos elaborar, entonces, la siguiente definición:

Si B es independiente de A también A será independiente de B. Por lo tanto, **los eventos A y B son independientes si se cumple:**

- P (B/A) = P (B/ $\overline{A}$) = P (B)
- P(A/B) = P(A/ $\overline{B}$) = P(A)

Cuando hablamos de probabilidad condicional, dijimos que $P(B/A) = \frac{P(A \cap B)}{P(A)}$

entonces, si A y B son independientes, o sea que P(B/A) = P(B) , se tiene:

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

De la ecuación anterior se deduce que: P(A ∩ B) = P(A) * P (B)

Luego, otra definición de independencia puede expresarse como:

Los eventos A y B son independientes si se cumple que:

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

Para el **Ejemplo 9**, verificaremos si el hecho de que una persona padezca dolores de cabeza es independiente del hecho de fumar. Si son independientes, debe cumplirse la siguiente igualdad:

$$P(A1 \cap B1) = P(A1)\ .\ P(B1)$$

$$P(A1 \cap B1) = \frac{12}{35} = 0{,}34$$

$$P(A1) = \frac{12}{21} \quad \text{y} \quad P(B1) = \frac{12}{20}$$

$$P(A1)\ .\ P(B1) = \frac{12}{21}\ .\ \frac{12}{20} = \frac{144}{420} = \frac{12}{35} = 0{,}34$$

Se cumple que $P(A1 \cap B1) = P(A1)\ .\ P(B1)$ Así concluimos que A1y B1son independientes.

De la misma manera, podemos comprobar la independencia de otros eventos, en el **Ejemplo 9.**

6. Teorema de Bayes

Antes de enunciar este teorema es necesario definir otros conceptos previos.

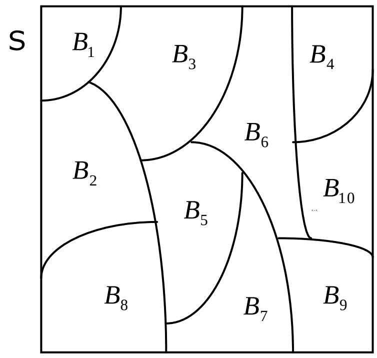
6.1 Partición

Decimos que los sucesos B1, B2, B3, …, Bk representan una partición del espacio muestral S si:

(a) $B_i \neq B_j$ para todo $i \neq j$

(b) $\bigcup_{i=1}^{k} B_i = S$ (S espacio muestral)

(c) $P(B_i) > 0$ para todo i

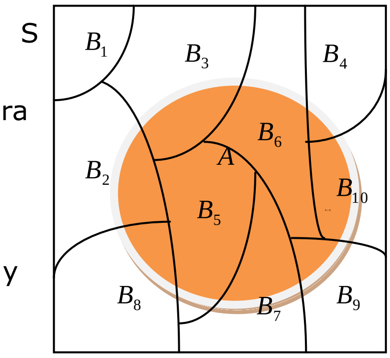


En otras palabras: cuando se efectúa el experimento E ocurre uno y solo uno de los sucesos B_i .

Ejemplo 10

En el lanzamiento de un dado $B_1 = \{1,2\}$, $B_2 = \{3,4,5\}$ y $B_3 = \{6\}$ representan una partición del espacio muestral, mientras que $C_1 = \{1,2,3,4\}$ y $C_2 = \{4,5,6\}$ no lo es.

Sea A algún suceso con respecto a S y sea B_1, B_2, …, B_k una partición de S, algunos de los conjuntos $A \cap B_i$ pueden ser vacíos, esto no invalida la anterior descomposición de S.



Por lo tanto, es posible aplicar la propiedad aditiva para este tipo de sucesos $P(A) = 1$ y escribir:

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_k)$$

Sin embargo, cada término $P(A \cap B_j) = P(A / B_j) \,.\, P(B_j)$ y por lo tanto, obtenemos el llamado *teorema de la probabilidad total:*

$$\boxed{P(A) = P(A / B_1) \,.\, P(B_1) + P(A / B_2) \,.\, P(B_2) + \ldots + P(A / B_k) \,.\, P(B_k)}$$

Este resultado representa una relación muy útil, ya que frecuentemente cuando se busca P(A) puede ser difícil calcularlo directamente. Sin embargo, con la información adicional de que B_j ha ocurrido, podemos calcular cada $P(A / B_j)$ y entonces usar la fórmula anterior.

6.2 Teorema de Bayes o Teorema de la Probabilidad de las causas

Como se dijo anteriormente, para una partición del espacio muestral, sucede uno y solo uno de los sucesos B_i.

Por lo tanto, para calcular la probabilidad de un B_i particular (esto es una "causa"), dado que el suceso A ha ocurrido, se aplica la definición de probabilidad condicional, se obtiene:

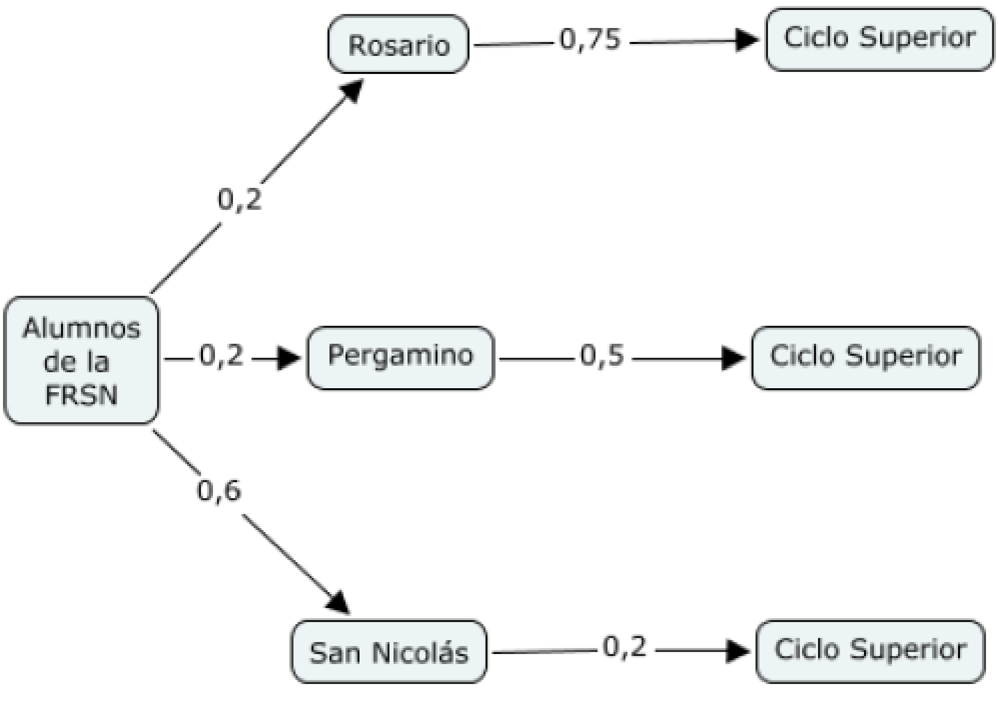
$$P(B_i / A) = \frac{P(A/B_i)\ .\ P(B_i)}{\sum_{j=1}^{k} P(A/B_j)\ .\ P(B_j)} \qquad i = 1,2,...,k$$

Este resultado se lo conoce como **Teorema de Bayes** o Teorema para la probabilidad de las "causas".

Ejemplo 11

El 20% de los alumnos de una Facultad son de Rosario y otro 20% son de Pergamino, el resto son de San Nicolás. El 75% de los alumnos de Rosario cursan el ciclo superior y el 50% de los de Pergamino también, mientras que de los de San Nicolás solamente el 20% cursan el ciclo superior. ¿Cuál es la probabilidad de que un alumno del ciclo superior elegido al azar sea de Rosario?

Un método muy útil para analizar este tipo de problemas es presentar la información a través de un diagrama de árbol:



Reemplazando en la fórmula del Teorema de Bayes:

$$P(Rosario / CicloSuperior) = \frac{0,2\ .\ 0,75}{0,2\ .\ 0,75 + 0,2\ .\ 0,5 + 0,6\ .\ 0,2} = 0,405$$

7. Algunas consideraciones finales sobre Probabilidad

Una manera de asignar probabilidad a un evento consiste en considerar la proporción de resultados favorables a éste respecto de todos los resultados posibles del experimento.

Esta definición de probabilidad se desprende de la definición clásica de probabilidad.

También, existen muchas situaciones en las que no se puede aplicar este procedimiento.

Por ejemplo, es dificultoso establecer la probabilidad de que mañana llueva utilizando este método. Una manera más satisfactoria de relacionar probabilidades al mundo real es en términos de frecuencia relativa.

Utilizando esta concepción teórica de la probabilidad, se establece que lanzando una moneda indefinidamente, y después de un cierto número de tiradas, se comprobaría que el cociente entre el número de veces que salió "cara" y el número total de veces que fue lanzada, sería muy próximo a cierto número. Y éste, será considerado como la probabilidad de obtener "cara".

Es decir, la probabilidad de un evento puede ser pensada como la frecuencia relativa con que éste ocurre, en un número indefinidamente grande de ensayos.

Aproximar la medida de la probabilidad a la frecuencia relativa, también tiene sus problemas. Por un lado la determinación es hipotética (se habla de tirar una moneda "indefinidamente").

Y por otra parte, la noción de que la frecuencia relativa se aproxima a un valor constante en un gran número de ensayos, necesita un estudio cuidadoso.

Sin embargo, al definir la probabilidad como el límite de la frecuencia relativa, establecemos una relación válida y estrecha con los experimentos reales.

Así, aunque existen otras teorías que definen de distintas maneras la noción de probabilidad, para los fines de la enseñanza de las probabilidades, bastará con atenernos a la definición clásica, (proporción de resultados favorables respecto de los posibles) y a la que establece la teoría frecuencial (límite de la frecuencia relativa).

En esta **Unidad Nº2** se han desarrollado definiciones y propiedades referidas a experimentos aleatorios y eventos. También las distintas teorías en que se basa el cálculo de las probabilidades asociadas a ellos.

Ahora bien, cuando asignamos números reales a los resultados de un experimento aleatorio, surge la definición de **variable aleatoria**. Este concepto es de fundamental importancia en la aplicación de la teoría estadística inferencial.