Probabilidades. Primera parte

Dra. Silvia María Pozo Abreu

La Habana, 2016

Este documento contiene las memorias del módulo Probabilidades perteneciente al primer año de la residencia de Bioestadística. Espero pueda serle útil a alguien más.

Son mis notas de clases, comentarios y revisión de la bibliografía sobre el tema.

Desde la posición de alumna de postgrado he podido percibir la manera en que la Bioestadística me resulta útil y necesaria. Con mis conocimientos podré transmitir nuevos hallazgos y conocimientos a estudiantes y profesionales de la salud en un lenguaje que puedan comprender.

Por ello los residentes debemos familiarizarnos con el lenguaje propio de la Bioestadística y así valorar de forma objetiva y crítica cualquier literatura científica.

Espero que este documento resulte de utilidad para residentes de Bioestadística y de cualquier otra especialidad médica que esté interesado en adentrarse en el mundo de la investigación científica.

Palabras clave: probabilidades, teoría de las probabilidades, variable aleatoria, bioestadística, distribución de probabilidad

1.1-INTRODUCCIÓN

Un objetivo de la Estadística inferencial es sacar conclusiones válidas para un conjunto de individuos u objetos (población) a partir del estudio de una parte de dicho conjunto (denominada muestra). Otro objetivo es el de medir el grado de incertidumbre presente en dichas inferencias en términos de probabilidades.

La teoría de las probabilidades constituye una herramienta fundamental en la Estadística inferencial ya que sirve de instrumento de medición del grado de incertidumbre presente en la inferencia; es un elemento de vital importancia, al intentar modelar matemáticamente las variables que se estudien y además por las aplicaciones prácticas inmediatas que se derivan de este propio concepto.

Es necesario tener en cuenta algunos aspectos fundamentales sobre la teoría de las probabilidades. Pasemos a verlos.

1.2-Probabilidad. Algunos aspectos fundamentales

La Teoría de las probabilidades es el estudio de modelos para experimentos aleatorios.

Un experimento es aquel que se realiza artificialmente bajo determinadas condiciones o se realiza independientemente de la voluntad del experimentador. Es un conjunto de condiciones y acciones determinadas bajo las cuales se observa un determinado fenómeno.

Todo experimento aleatorio tiene tres componentes: condiciones, acción y observación.

Existen dos tipos de experimentos: el determinista y el aleatorio.

El experimento determinista es aquel que bajo un conjunto de condiciones o causas, el resultado del experimento queda perfectamente determinado, por ejemplo, solidificación del agua.

El experimento aleatorio es aquel que bajo las mismas condiciones puede dar lugar a diferentes resultados o efectos del experimento, por ejemplo, lanzamiento de un dado, peso de un recién nacido.

La palabra aleatorio significa casual, al azar.

Sucesos o eventos aleatorios

Cuando se realiza un experimento aleatorio diversos resultados son posibles. Es decir, todo experimento aleatorio tiene asociado un espacio muestral definido.

Espacio muestral: es el conjunto de todos los posibles resultados de un experimento aleatorio. Se le denota por la letra S.

El conjunto base está formado por N elementos; el tamaño de la muestra es n.

1. Seleccionar aleatoriamente n elementos del conjunto N elementos todos a la vez:

El espacio muestral tiene N(S): cantidad de elementos de S (es el tamaño del espacio muestral):

N(S)= $\binom{N}{n}$ donde $\binom{N}{n}$ significa combinaciones de N en n y es el coeficiente binomial o número combinatorio.

Entonces $\binom{N}{n}= \frac{N!}{n!(N-n)!}$ donde N! es el factorial de N y N!=1.2.3…N.

2. Seleccionar aleatoriamente una muestra de tamaño n uno a uno sin reposición:

N(S)= $V_{N,n}$= V_n^N donde V_n^N son las variaciones de N en n.

$$V_n^N = \frac{N!}{n!\,(N-n)!}$$

3. Seleccionar aleatoriamente una muestra de tamaño n, uno a uno con reposición:

N(S)= N^n

Cuando no hay orden es lo mismo que decir todos a la vez y si hay orden quiere decir uno a uno.

Suceso o evento: es cualquier característica observada del resultado de un experimento. Es un subconjunto del espacio muestral que lo denotaremos con letras mayúsculas del alfabeto (A, B, C…).Cada uno de los elementos de un espacio muestral es un suceso elemental.

El espacio muestral y un suceso se pueden representar en forma de conjunto o en forma gráfica.

Probabilidad: la probabilidad de un evento A, es un número real en el intervalo [0, 1] que se denota por P(A), y representa una medida de la frecuencia con la que se observa la ocurrencia del evento A cuando se efectúa el experimento aleatorio en cuestión.

La probabilidad de un suceso es una medida cuantitativa de la posibilidad de ocurrencia de ese suceso.

Hay diferentes tipos de sucesos:
-ciertos: los que siempre ocurren (posibilidad de ocurrencia 100 %)
-imposibles: nunca ocurren posibilidad de ocurrencia 0 %)
-aleatorios: pueden o no ocurrir (posibilidad de ocurrencia $0 \leq x \leq 100$ %)

La intervención es un modo de cambiar las condiciones de un suceso aleatorio.

Existen varias definiciones de probabilidad:

Definición clásica de probabilidad:

$$P(A) = \frac{N(A)}{N(S)}$$

Se dice "N de A". Es la cantidad de sucesos favorables o cantidad de eventos o elementos de A

Se dice "N de S". Es el número de eventos posibles o cantidad de sucesos posibles

La probabilidad del evento A está entre cero y uno:
$0 \leq P(A) \leq 1$

Esta definición es únicamente válida para espacios muestrales finitos.

<u>Definición frecuentista de probabilidad:</u>
También conocida como probabilidad estimada o empírica de un evento.

Es la frecuencia relativa con que se presentará un suceso si un fenómeno aleatorio se repite indefinidamente.

Formalmente:

Suponga que se realizan n repeticiones de un cierto experimento aleatorio y sea A un evento cualquiera. Denotemos por n(A) el número de ocurrencias del evento A, en las n realizaciones del experimento. Se define entonces la probabilidad frecuentista de A como indica el siguiente límite:

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

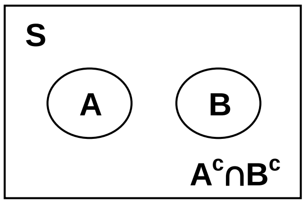
Esta definición estadística de probabilidad tiene dificultades matemáticas (aunque es muy útil en la práctica) ya que puede ser que no exista un verdadero número límite. Debido a ello la teoría de probabilidad moderna ha sido desarrollada en forma axiomática, es decir el concepto de probabilidad se deja sin definir.

Pasemos ver el concepto de probabilidad axiomática.

Probabilidad axiomática:
La definición axiomática de la probabilidad no establece la forma explícita de calcular las probabilidades sino únicamente se proponen las reglas que el cálculo de probabilidades debe satisfacer.

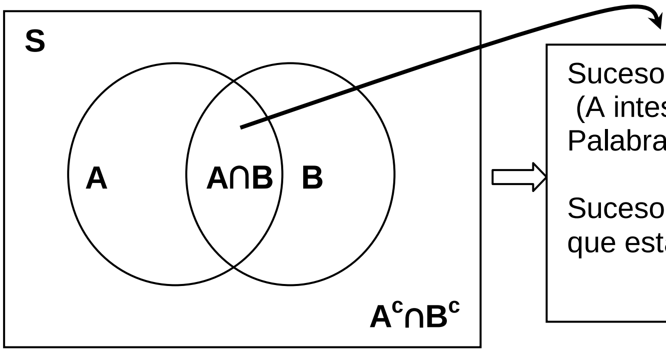
Axiomas de la probabilidad:

1) $P(A) \geq 0$.
2) $P(S) = 1$.
3) $P(A \cup B) = P(A) + P(B)$, cuando $A \cap B = \emptyset$. Esto es: Disjuntos (excluyentes)



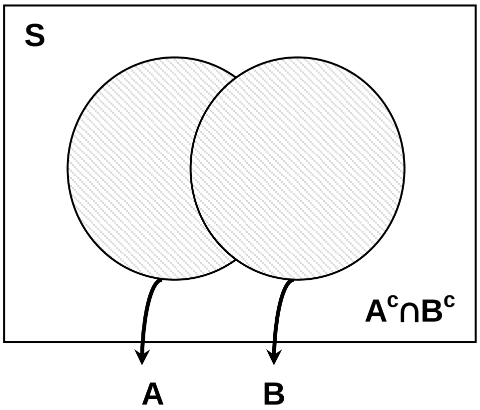
1.3- Teoría de conjuntos

Es necesario recordar la para poderla insertar aquí con las probabilidades.



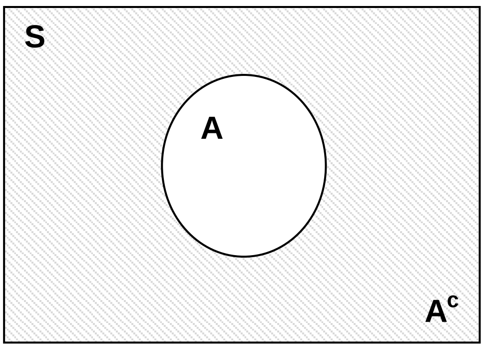
Suceso intersección $A \cap B$
(A intesección de B)
Palabra clave "y"

Suceso formado por los elementos que están en A y B simultáneamente.



Unión $A \cup B$
Palabra clave "o"

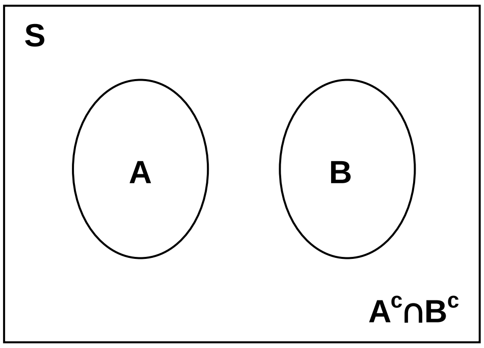
Suceso formado por los resultados experimentales que están en A o en B (incluyendo los que están en ambos).



Complemento A^c

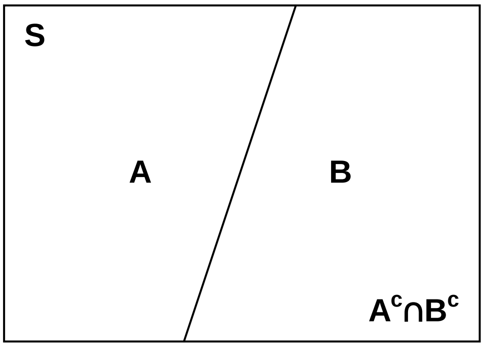
Palabra clave "no"

Conjunto de los elementos de U que no pertenecen a A



Mutuamente excluyentes (Disjuntos)
P(A∩B)= ∅
A∩B= ∅

No pueden ocurrir en forma simultánea, es decir si y sólo si su intersección es vacía



Exhaustivos (es el espacio muestral)

P(A∪B)= 1

A∪B= S

Los diagramas de Venn en ningún momento constituyen una prueba matemática, sin embargo, permiten tener una visión intuitiva de la relación que puede existir entre los conjuntos. Ver también el Anexo 1.

1.4-Propiedades de la probabilidad

a) $P(A^c) = 1 - P(A)$

b) $P(\emptyset) = 0$

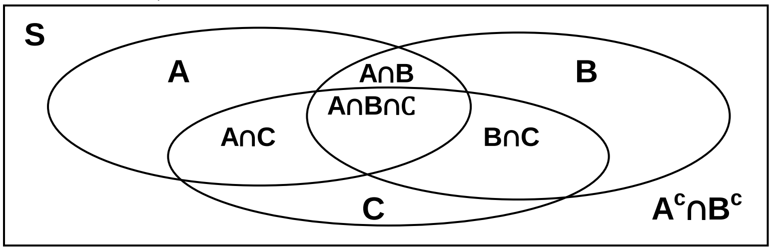
c) Si $A \subseteq B$, entonces $P(A) \leq P(B)$

d) Si $A \subseteq B$, entonces $P(B - A) = P(B) - P(A)$

e) $0 \leq P(A) \leq 1$

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

g) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C)$
$-P(B \cap C) + P(A \cap B \cap C)$.



2-Fundamentos de la teoría de probabilidades

2.1-Probabilidad condicional. Independencia de sucesos. Fórmula de probabilidad completa. Teorema de Bayes. Independencia de sucesos.

Sucesos o eventos mutuamente excluyentes: cuando la ocurrencia de uno excluye (evita) la ocurrencia del otro y se cumple $P(A \cap B) = 0$.

Para sucesos mutuamente excluyentes se cumple: P(A+B)= P(A) + P (B), que es lo mismo que decir $P(A \cup B) = P(A) + P(B)$ (y es lo mismo que decir P(A o B)= P(A) + P(B))

Eventos independientes (independencia de sucesos): cuando la ocurrencia de uno de ellos no afecta en lo más mínimo la ocurrencia del otro.

En el caso de suceso independiente se define que:

$P(A \cap B) = P(A) \, . \, P(B) \Longrightarrow$ Regla del producto de las probabilidades

Probabilidad condicional: la probabilidad de ciertos eventos depende o se ve influenciada por la ocurrencia de otros. Formalmente:

$P(A/B)= \frac{P(A \cap B)}{P(B)}$ Palabras clave: si, dado que, sabiendo que.

Se denomina "Probabilidad de A dado que".

Cuando dos eventos no son independientes, sino que están relacionados de alguna manera, se utiliza el concepto de eventos condicionados o relacionados. Es decir, si la ocurrencia de uno de ellos condiciona de alguna manera la ocurrencia del otro se dice que los eventos están relacionados.

Es decir, si A y B son independientes:

$$\left\{ \begin{array}{l} P(A/B)= P(A) \\ P(B/A)= P(B) \end{array} \right. \Longrightarrow$$

Esto es así porque: $P(A/B)= \frac{P(A \cap B)}{P(B)} = P(A)$

Esta es la definición, donde

$P(A \cap B)= P(A) \, . \, P(B)$ (entonces A y B son sucesos independientes

Teorema de la multiplicación de probabilidad:

Sea P una probabilidad y $A_1, A_2, \ldots A_n$ sucesos tales que $P(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$ la probabilidad de intersección de los n-1 primeros sea mayor que cero, entonces

$$P(A_1 \cap A_2 \cap \ldots \cap A_n)= P(A_1).P(A_2/A_1).P(A_3/A_1 \cap A_2)\ldots$$

$$\boxed{P(A_n/A_1 \cap A_{2 \ldots} \cap A_{n-1})}$$

Un error frecuente consiste en confundir la probabilidad condicional con la probabilidad de la intersección: en ambas medimos efectivamente la intersección, pero…

-En $P(A \cap B)$ con respecto a P(S)=1

-En P(A|B) con respecto a P(B)

"tamaño" de uno respecto al otro

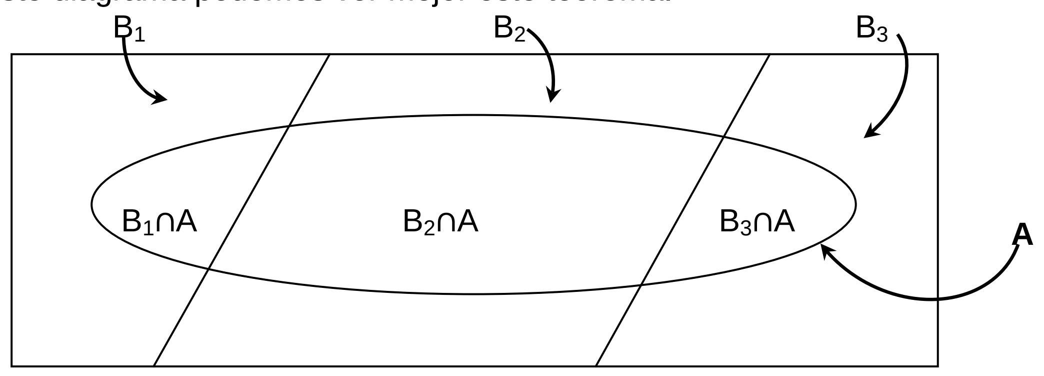
El concepto de independencia se hace extensivo a las variables aleatorias y muchas veces nos enfrentaremos al problema de comprobar si dos resultados posibles o dos variables aleatorias son independientes. Por ejemplo, cuando investigamos sobre posibles factores de riesgo para alguna enfermedad o condición que influye sobre la salud de un individuo lo que se busca, en sentido general, son los resultados o variables que están asociados a la enfermedad o condición, y son previos a esta última: se trata de probar la independencia.

Teorema de la probabilidad total:
Aquí el especio muestral es una unión disjunta de sucesos con probabilidades mayores que cero.

Definición del Teorema de la probabilidad total: si el espacio muestral de un experimento aleatorio se puede expresar como la unión disjunta de los sucesos A_1, A_2, … A_n los cuales tienen todos probabilidades mayores que cero, entonces se cumple que la probabilidad de A es:
$P(A)= P(A_1).P(A/A_1) + P(A_2).P(A/A_2) + \ldots + P(A_n).P(A/A_n)$

$$\boxed{P(A)= \sum_{i=1}^{n} P(A_i).P(A/A_i)}$$

En este diagrama podemos ver mejor este teorema.



$$P(A)= \overbrace{P(B_1) \,.\, P(A/B_1)}^{P(B_1\cap A)} + \overbrace{P(B_2) \,.\, P(A/B_2)}^{P(B_2\cap A)} + \overbrace{P(B_3) \,.\, P(A/B_3)}^{P(B_3\cap A)}$$

donde P(A) es la probabilidad total de A.
Con esa probabilidad total calculada se puede calcular el Teorema de Bayes.
Fíjense que he repetido la probabilidad total para que se vea su relación:

$$P(A)= \overbrace{\underline{P(B_1) \,.\, P(A/B_1)}}^{P(B_1\cap A)} + \overbrace{\underline{P(B_2) \,.\, P(A/B_2)}}^{P(B_2\cap A)} + \overbrace{\underline{P(B_3) \,.\, P(A/B_3)}}^{P(B_3\cap A)}$$

Teorema de Bayes:

$$P(B_1/A)= \frac{P(B1\cap A)}{P(A)} = \frac{P(B1)\,.P(A/B1)}{P(A)}$$

$$P(B_2/A)= \frac{P(B2\cap A)}{P(A)} = \frac{P(B2)\,.P(A/B2)}{P(A)}$$

$$P(B_3/A)= \frac{P(B3\cap A)}{P(A)} = \frac{P(B3)\,.P(A/B3)}{P(A)}$$

A este teorema se le llama también probabilidad de los sucesos o probabilidad a posteriori.

Definición formal del teorema de Bayes: bajo los supuestos del teorema de la probabilidad total, se cumple que:

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

Este teorema tiene una aplicación básica en Medicina para decidir si un paciente padece cierta enfermedad o no, en función de los resultados de un test diagnóstico.

A este teorema se le llama también probabilidad de los sucesos o probabilidad a posteriori.

Algunas palabras clave para ubicarnos en lo que nos piden:
Cuando dicen “al menos una vez” es igual a decir “alguna vez se obtiene”, por lo tanto su complemento es “ninguna vez se obtiene”.

Cuando hay reposición los eventos son independientes ya que cuando se repone algo la ocurrencia de un evento no afecta a otro. Es decir: muestras con reposición = muestras independientes.

Cuando no hay reposición los eventos son dependientes ya que la ocurrencia de uno sí afecta la ocurrencia de otro.

Ejemplos del uso de la terminología de probabilidad en la medicina:
- Un paciente con determinados signos y síntomas se dice tiene una posibilidad del 80% de tener la enfermedad x.
- La posibilidad de un paciente de sobrevivir a una operación urgente de apendicitis es alta, más del 98%.
- Prevalencia de asma bronquial en el municipio Playa perteneciente a la ciudad de La Habana es del 10%.

Ahora están en condiciones de ir al anexo 2 y ver ejemplos de cálculo de probabilidades.

También pueden ir al anexo 3 y ver cómo son las probabilidades en una tabla de contingencia y diagrama de árbol.

En el anexo 4 podrán ver mejor algunas aplicaciones de la teoría de la probabilidad en la práctica médica.

Para recordar la Ley de la regularidad estadística, ya estudiada durante nuestros años en pregrado pueden leer el anexo 5.

2.2-Variables aleatorias. Definición. Variable aleatoria discreta y continua. Independencia de variables aleatorias

Variable aleatoria
Definición: es una función que atribuye un único número real a cada suceso elemental del espacio muestral del experimento aleatorio.

Se denota como: $X : S \rightarrow \mathbb{R}$
Es decir transforma en números los resultados y los probabiliza.

Debe quedar claro lo qué es una variable aleatoria en la práctica, es decir, interpretar esa definición de lo que es la variable aleatoria. En el Anexo 6 pueden estudiar mejor la variable aleatoria, pero no dejen de verlo porque es un aspecto básico en la Bioestadística su comprensión.

A pesar de haber leído el anexo indicado vuelvo a recalcar con otras palabras que la **variable aleatoria es aquella sobre la que actúan factores no controlables o desconocidos (azar), los cuales provocan la imposibilidad de predecir el valor que tomará; son producto de un suceso aleatorio.** Por lo que les pido que lean y relean interiorizando lo aquí escrito en relación con la variable aleatoria, uno de los conceptos fundamentales en la Bioestadística que deben dominar desde ahora y que les va a acompañar a partir de aquí.

Estas variables aleatorias (v.a) pueden ser:
-v.a. discreta: es aquella que sólo puede tomar un número finito **numerable** de valores. Ejemplo: sexo de un recién nacido.
X= 0, 1, 2, …, n (v.a. discreta) $\longrightarrow$ El conjunto de valores que ésta (la variable) toma es un conjunto discreto (numerable).

-v.a. continua es la que puede tomar un número infinito **no numerable** de valores. Ejemplo: peso de un recién nacido.

X= (a,b) ≤ $\mathbb{R}$ (v.a. continua) $\longrightarrow$ Cuando toma todos los valores dentro de un intervalo (a,b) ⊆ $\mathbb{R}$.

2.3-Función de distribución. Función de distribución univariada. Definición y propiedades

Las variables aleatorias, transforman eventos del espacio muestral en eventos numéricos, los cuales tienen asociada una probabilidad de ocurrencia.

Distribución de probabilidad de una variable aleatoria

Es una tabla, gráfica o expresión matemática (fórmula) que da las probabilidades con que una variable aleatoria toma diferentes valores.

La tabla de distribución:

X	X_1	X_2	X_3	X_4
P	P_1	P_2	P_3	P_4

Aquí Pk= P(X=Xk). Por ejemplo, en la tabla anterior veremos:
P1= P(X=X1)
P2= P(X=2) hasta llegar a P4.
Permite caracterizar una variable aleatoria discreta

Funciones de probabilidad

Son dos; una llamada función de probabilidad (es para v.a. discretas) y otra llamada función de densidad (para v.a. continuas). Pasemos a describirlas.

Función de probabilidad de una v.a. discreta: es el conjunto de pares ordenados (X, $f(x)$) donde X es el conjunto de valores de una variable aleatoria y $f(x)$ las probabilidades asignadas a X. Esta función tiene la siguiente forma:

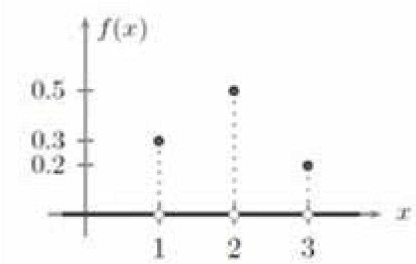
$$f(x) = \begin{cases} P(X = x) & \text{si } x = x1,\ x2,\ \ldots \\ 0 & \text{otro caso.} \end{cases}$$

Se cumplen las dos propiedades:

a) $f(x) \geq 0, \quad \forall\, x \in \mathbb{R}$

El símbolo ∀ significa "para todo"

b) $\sum_x f(x) = 1.$



x	1	2	3
$f(x)$	0.3	0.5	0.2

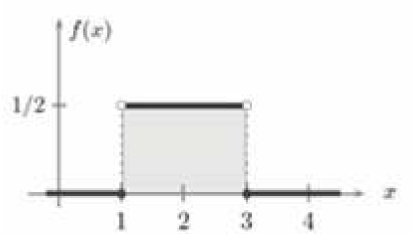
Para que se comprenda mejor, esta función de probabilidad es el conjunto de todos los valores de la variable junto con sus probabilidades.

Función de densidad de una v.a. continua: así se le llama a la función de probabilidad de una variable aleatoria continua. Es aquella que cumple las siguientes propiedades:

a) $f(x) \geq 0, \quad \forall\, x \in \mathbb{R}$

El símbolo ∀ significa "para todo"

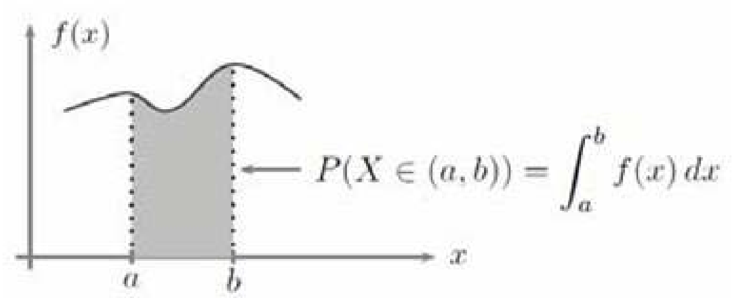
b) $\int_{-\infty}^{\infty} f(x)dx = 1$



Es una función no negativa, cuya integral vale 1, como ya vimos.

Aquí la probabilidad de que la variable tome un valor dentro del intervalo (a, b) se puede calcular o expresar como el área bajo la función de densidad en el intervalo (a, b). De esta forma el cálculo de una probabilidad se reduce al cálculo de una integral.

$P(a < X < b) = P(a \leq X \leq b) = \int_a^b f(x)dx$



Función de distribución (acumulada)

Para toda variable aleatoria, ya sea discreta o continua, la función de distribución evaluada en un número x cualquiera es simplemente la probabilidad de que la variable aleatoria tome un valor menor o igual a x, o en otras palabras, que tome un valor en el intervalo $(-\infty, x]$ Se denota por F(x).

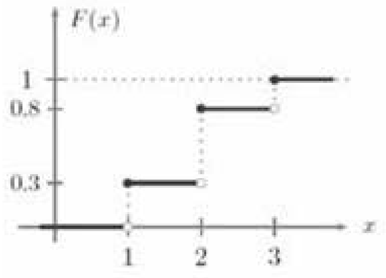
$$F(x) = P(X \leq x)$$

Sus valores están siempre entre 0 y 1.

Esta función se puede calcular a partir de la función de probabilidad o densidad.

En caso de v.a. discreta: $F(x) = P(X \leq x) = \sum_{\mu < x} P(X = \mu)$

Su gráfico es el siguiente:



Como puede observarse en la gráfica anterior, para una variable aleatoria discreta, la función de distribución acumulada es una función escalonada, una función que solo se incrementa por saltos finitos y es constante entre dos valores adyacentes cualesquiera de la variable.

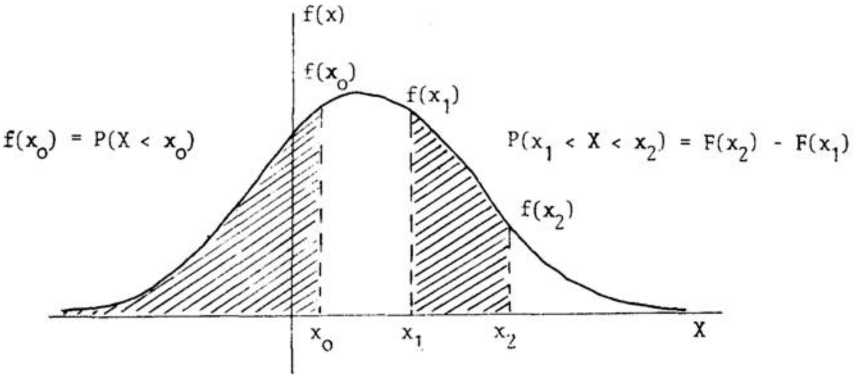
En caso de v.a. continua: $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx$

Propiedades de la función de distribución (acumulada):

1. $0 \leq F(x) \leq 1$
2. $\lim_{x \to \infty} F(x) = 1$
3. $\lim_{x \to -\infty} F(x) = 0$
4. Si $x_1 \leq x_2$ entonces $F(x_1) \leq F(x_2)$
5. Si $x_1 \leq x_2$ entonces $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$

Al ser la variable aleatoria continua, tiene un número incontable de valores posibles. La función de distribución acumulada será continua en el sentido de que su gráfica es suave, es decir, no hay saltos en la probabilidad en cualquiera de los valores de la variable. Debido a la continuidad, la probabilidad de X en un punto cualquiera de sus valores posibles **es cero**. Así, aunque la función de distribución acumulada de una variable aleatoria continua puede ser aplicada, no da resultados útiles. En su lugar, representamos la función de X, f(x) como una función densidad de probabilidades. Por tanto, la función densidad es una medida de la concentración de probabilidad dentro de un intervalo. Esta probabilidad puede interpretarse como un área (una integral) bajo la curva de f(x), llamada **curva de densidad,** limitada por las ordenadas de los extremos del intervalo. Más, precisamente, si X es una variable aleatoria continua, la probabilidad de que asuma un valor en el intervalo (x_1, x_2), puede representarse por $P(x_1 < X < x_2)$ y es igual al área bajo la curva de f(x) encerrada por las ordenadas erigidas en x_1 y x_2. Esto es importante que o entiendan.

Su gráfica es:



2.4-Esperanza matemática y varianza de una variable aleatoria

Esperanza: es un número que indica el promedio de los diferentes valores que puede tomar la variable. También se le llama valor esperado, valor promedio, media y promedio. Se denota por la letra griega μ (mu).

En caso de las v.a. discretas: $\mu = E(X) = \sum_x xP(X = x) = \sum_x x\, f(x)$

En caso de las v.a. continuas: $\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)dx$

Varianza: es la medida del grado de dispersión de los diferentes valores tomados por la variable. Se denota por Var(X) y también por la letra griega σ^2 (sigma cuadrada).

Para calcular la varianza se necesita conocer primero la esperanza:

En caso de las v.a. discretas: $Var(X) = \sum_x (x - E(X))^2 f(x)$

En caso de las v.a. continuas: $Var(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx$

A la raíz cuadrada positiva de la varianza (σ) se le denomina desviación estándar: $\sigma = \sqrt{Var(X)}$

A mayor desviación estándar mayor es la diferencia entre los valores y su promedio, es decir, más disperso están los datos.

Hasta aquí la primera parte sobre Probabilidades.

Es necesario que comprendan todo porque esta es una parte básica de la Bioestadística. Lo aquí estudiado se retomará en módulos posteriores como por ejemplo: Muestreo y Estadística inferencial.

En la segunda parte estudiaremos las distribuciones de probabilidad más utilizadas, haciendo énfasis en la distribución normal. Les adelantaré que el comportamiento en la población de las variables aleatorias puede ser descrito usando un modelo matemático que brinde una expresión analítica del fenómeno, este es el caso de los modelos teóricos de probabilidad.

Se anexa un glosario que contiene algunos términos en español e inglés que son de uso en este módulo. Ver anexo 7.

3-Bibliografía consultada

1. Massip Nicot J. Probabilidades. Conferencia impartida a residentes de primer año de Bioestadística. INHEM: La Habana, 2015.
2. Pérez Piñero J. Inferencia estadística. Generalidades. Conferencia impartida a residentes de segundo año de la especialidad de Bioestadística. La Habana: ENSAP; 2015.
3. Bayarre Vea H, Oliva Pérez M. Introducción a la Estadística Inferencial, en: Métodos y Técnicas Aplicados a la Investigación en Atención Primaria de Salud. Teoría y Práctica. Ciudad de La Habana: ISCM-La Habana; 2004.
4. Bayarre Vea H. Estimación puntual y por intervalos de confianza. Conferencia impartida a residentes de segundo año de la especialidad de Bioestadística. La Habana: ENSAP; 2015.
5. Bayarre Vea H. Pruebas de hipótesis. Conferencia impartida a residentes de segundo año de la especialidad de Bioestadística. La Habana: ENSAP; 2015.
6. Callaert H. Concepts of Probability and Statistics.
7. Rosner B. Fundamentals of Biostatistics. 5ta ed. Duxbury; 2000.
8. Rubén Quesada M. Probabilidad. Bibliografía básica de la signatura Inferencia II para estudiantes de licenciatura en Gestión de la información en salud. CECAM: La Habana, 2009.
9. Rubén Quesada M. Inferencia estadística. Conferencia impartida a residentes de segundo año de la especialidad de Bioestadística. La Habana: INHEM; 2014.
10. MINSAP. Bioestadística y computación. Ed. Pueblo y Educación, La Habana, 1987.
11. Camel Vargas F. Estadísticas Médicas y de Salud Pública.Editorial Pueblo y Educación: La Habana, 1968.
12. Colectivo de autores. Capítulo 9. Teoría de las probabilidades: nociones básicas. Introducción a la Inferencia Estadística. Informática Médica. Tomo 2. Bioestadística. Centro de Cibernética Aplicada a la Medicina (CECAM): La Habana; 2004.
13. Cruz Pérez NR: Bioestadística. Ed. Universo Sur: Cienfuegos; 2013.

14. Barón López FJ, Rius Díaz F. Bioestadística. Paraninfo; 2008. ISBN: 978-8497323413. Disponible en: http://www.bioestadistica.uma.es/baron/apuntes/
15. Ruiz-Maya Pérez L, Martín Pliego J. Fundamentos de inferencia estadística. Madrid: Editorial AC; 1999.
16. Martínez González MA, Sánchez Villena A, Faulin Fajardo J. Bioestadística amigable. 2da edición. Editorial: Díaz de Santos; 2006.
17. Esper R. J, Machado R.A. La investigación en medicina. Bases teóricas y prácticas. Elementos de bioestadística. Ghione Impresores; 2008.
18. Díaz Narváez V.P. Metodología de la investigación científica y bioestadística. Ril Editores; 2006.
19. Rincón L. Curso elemental de Probabilidad y Estadística. UNAM: México DF; 2007. Disponible en: http://www.matematicas.unam.mx/lars
20. Spiegel MR, Stephens LJ. Estadística. 4ta edición. Ed. McGRAW-HIL: México D.F; 2009.
21. Joaquín Seda J. Probabilidad. Bibliografía complementaria del módulo Probabilidades de La residência de Bioestadística. INHEM: La Habana; 2014.
22. Marques de Cantú M.J. Probabilidad y Estadística para Ciencias Químico Biológicas. McGRAW-HILL; 1991.

ANEXO 1. Teoría de conjuntos. Algunos aspectos

-Unión $\longrightarrow$ $A \cup B$

Todos los elementos que están en A y B.

(Es conmutativa)

-Intersección $\longrightarrow$ $A \cap B$

Elementos que están en los 2.

(Es conmutativa)

-Complemento $\longrightarrow$ A^c

Lo que le falta al conjunto A para ser el universo.

-Diferencia $\longrightarrow$ $A \setminus B$ (lo que está en A y no en B)

$B \setminus A$ (lo que está en B y no en A)

-Diferencia simétrica $\longrightarrow$ $(A \setminus B) \cup (B \setminus A) \equiv (A \cup B) \setminus (B \cap A) = A \vartriangle B$

(Es conmutativa)

$A \vartriangle B = \emptyset$ si y sólo si $A=B$

Leyes de Morgan:

$(A \cup B)^C = A^C \cap B^C$

$(A \cap B)^C = A^C \cup B^C$

ANEXO 2. Ejemplo de cálculos de probabilidades

Ejercicio 1. Se ha repetido en 1000 ocasiones el experimento de elegir a una mujer de una población muy grande y clasificar con respecto a las variables que se presentan en la siguiente tabla.

	Menopausia		Total
	No	Sí	
Sin afección ósea	189	280	469
Osteomielitis	108	359	467
Osteoporosis	6	58	64
Total	303	697	1000

a) ¿Cuál es la probabilidad de que una mujer tenga osteoporosis? Respuesta:

P(osteoporosis)= $\frac{64}{1000}$ = 0,064. Si se multiplica por 100 quedaría que la probabilidad de que una mujer tenga osteoporosis es 6,4%.
Esta es la noción frecuentista de probabilidad.

b) ¿Cuál es la probabilidad de tener osteomielitis u osteoporosis? Respuesta: Los sucesos osteomielitis y osteoporosis son disjuntos (osteomielitis ∩osteoporosis=Ø), entonces:

P(OsteomielitisUOsteoporosis)= $\frac{467}{1000} + \frac{64}{1000} = 0{,}531$

c) ¿Cuál es la probabilidad de tener osteoporosis o menopausia? Respuesta: ¿Qué sucede en este caso? Que osteoporosis y menopausia no son sucesos disjuntos luego:

P(osteoporosisUmenopausia)= $\frac{64}{1000} + \frac{697}{1000} - \frac{58}{1000} =$ 0,703

d) ¿Cuál es la probabilidad de una mujer sin afección ósea (SAO)?

Respuesta:

P(SAO)= $\frac{469}{1000} =$ 0,469

Pero también se pudo haber resuelto así:
P(SAO)= 1 - P(SAO^c)= 1 - P(osteomielitisUosteoporosis) = 1 - 0,531 = 0,469

e) Si es menopáusica ¿Cuál es la probabilidad de tener osteoporosis?

P(osteoporosis/menopausia)= $\frac{58}{697} =$ 0,098; esto es así porque:

P(osteoporosis/menopausia)= $\frac{58}{1000} / \frac{697}{1000}$ = 0,098

f) ¿Cuál es la probabilidad de tener menopausia y osteoporosis? Respuesta:

P(menop ∩osteoporosis) = $\frac{58}{1000}$ = 0,058

Concepto de independencia:

La idea de independencia es un concepto esencial en estadística. La independencia entre dos variables implica que la información que poseamos sobre una variable no sirve en absoluto para predecir la otra.

Veamos un ejemplo:

Se realizó una encuesta a 300 personas, 100 mujeres y 200 hombres y se les preguntó si son fumadores activos. Los resultados se muestran en la siguiente tabla.

Tabla 1. Hábito de fumar en una muestra según sexo.

	Fuman	No fuman	**Total**
Mujeres	20	80	**100**
Hombres	40	160	**200**
Total	**60**	**240**	**300**

La probabilidad de encontrar un fumador sería del 20%:

$$\text{P(Fumar)} = \frac{60}{300} = 0{,}2.$$

Esta probabilidad valdría lo mismo para los hombres que para las mujeres. Como esta probabilidad es la misma para hombres y para mujeres diremos que la probabilidad de fumar es independiente del sexo. Es decir, saber el sexo de alguien de la muestra (Tabla 1) no proporciona ninguna información que permita averiguar si hay mayor o menor probabilidad de que esa persona sea fumadora.

Otro ejercicio. Este es sobre probabilidad condicional.

Probabilidad condicionada: es el cociente entre los casos favorables y los casos posibles dentro de aquellos que cumplen una condición. Es la probabilidad de ocurrencia de un fenómeno dentro de un subgrupo. Se expresa con una barra vertical (/) como la notación que indica cuál es la condición.

Ejemplo:
Se está valorando la presencia de glucosa en orina (glucosuria) en pacientes diabéticos y en pacientes sin diabetes obteniendo la siguiente Tabla de contingencia:
Tabla 2. Presencia de glucosuria en pacientes diabéticos y no diabéticos.

	Diabetes	No diabetes	**Total**
Glucosuria	60	8	**68**
No glucosuria	140	792	**932**
Total	**200**	**800**	**1000**

Vemos que se han examinado 100 pacientes, 200 de ellos eran diabéticos y 800 no diabéticos.
Ahora la probabilidad de que un diabético presente glucosuria (probabilidad de glucosuria condicional a la diabetes) es:

$$\text{P(glucosuria/diabetes)}=\frac{60}{200} = 0{,}3$$

Esto es lo mismo que decir:

P(glucosuria/diabetes)=$\frac{casos\ con\ glucosuria\ y\ diabetes}{diabéticos} = \frac{P(gucosuria \cap diabetes)}{P(diabéticos)}$

Si este resultado se multiplica por 100 decimos que “el 30 % de los diabéticos presentan glucosuria”.

Ahora la probabilidad de que un no diabético presente glucosuria (probabilidad de glucosuria condicional a la diabetes) es:

P(glucosuria/no diabetes)=$\frac{8}{800} = 0{,}01$

Si este resultado se multiplica por 100 decimos que “sólo el 1 % de los no diabéticos presentan glucosuria”.

En general la probabilidad condicional se calcula como sigue:

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

Lo opuesto a la probabilidad condicionada es la probabilidad marginal (o no condicionada o probabilidad total). La probabilidad marginal de presentar glucosuria es:

P(glucosuria) = $\frac{casos\ con\ glucosuria}{total\ de\ pacientes}$= $\frac{68}{1000} = 0{,}068$

Volvamos a la tabla 1 (hombres y mujeres fumadores): teóricamente podríamos distinguir entre la probabilidad marginal de ser fumador que es la que se calcula teniendo en cuenta al rotal de los sujetos y que valdría:

Probabilidad de fumar: $P_{marginal}$= $\frac{60}{300} = 0{,}2$

y la la probabilidad de fumar *condicionada* a ser mujer, que sólo se refiere al subgupo

de mujeres y valdría:

Probabilidad de fumar: $P_{condicionada\ a\ ser\ mujer}$= $\frac{20}{100} = 0{,}2$

O si consideramos la probabilidad de fumar condicionada a ser hombre:

Probabilidad de fumar: $P_{condicionada\ a\ ser\ hombre}$= $\frac{40}{200} = 0{,}2$

Así comprobamos que la probabilidad marginal de ser fumador es idéntica a las probabilidades condicionales. **Esto es lo que sucede cuando hay independencia. Sólo si hay independencia coinciden las probabilidades condicionales y marginales.**

ANEXO 3. Tablas de contingencia, diagramas de árbol y cálculo de probabilidades

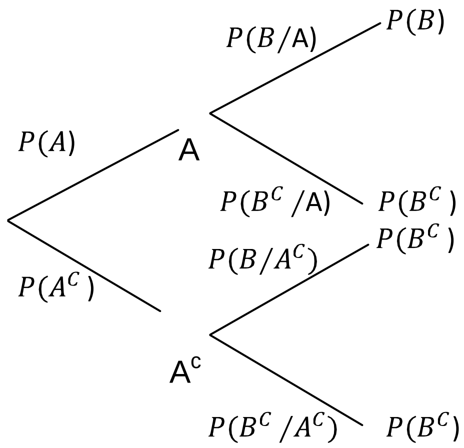
En los problemas de probabilidad y en especial en los de probabilidad condicionada, resulta práctico organizar la información en una tabla de contingencia o en un diagrama de árbol. Las tablas de contingencia y los diagramas de árbol están íntimamente relacionados, dado uno de ellos podemos construir el otro. Unas veces, los datos del problema permiten construir fácilmente uno de ellos y a partir de él podemos construir el otro, que nos ayudará en la resolución del problema.

Conversión de una tabla en diagrama de árbol
Las tablas de contingencia están referidas a dos características que presentan cada una dos o más sucesos.

	A	A^c	Total
B	$P(A \cap B)$	$P(A^c \cap B)$	$P(B)$
B^c	$P(A \cap B^c)$	$P(A^c \cap B^c)$	$P(B^c)$
Total	$P(A)$	$P(A^c)$	1

En el caso de los sucesos A, A^c, B y B^c expresados en frecuencias absolutas, relativas o probabilidades la tabla, adopta la forma adjunta.

Dicha tabla adopta la forma del diagrama de árbol del dibujo. En éste, a cada uno de los sucesos A, A^c se les ha asociado los sucesos B y B^c.



Sobre las ramas del diagrama de árbol se han anotado las probabilidades condicionadas correspondientes, deducidas de las relaciones análogas a:

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

Conversión de un diagrama en tabla de contingencia
De manera recíproca, dado el diagrama de árbol podemos construir la tabla de contingencia equivalente si más que utilizar la expresión:

$P(B \cap A) = P(B/A)\ .\ P(A)$, para calcular las probabilidades de las intersecciones de sucesos que forman la tabla.

ANEXO 4. Algunas aplicaciones de la teoría de la probabilidad en la práctica médica

Se puede afirmar que es muy amplio el empleo de la probabilidad en la práctica médica, por ejemplo, el médico al establecer sus diagnósticos basa su juicio en muchas ocasiones en razonamientos que en su esencia son probabilísticos, pues toma en cuenta, al hacerlo, toda una serie de experiencias análogas ya conocidas, que sirven de sustento al nuevo diagnóstico.

Otros ejemplos que pueden extraerse de los textos de clínica también están relacionados con el uso de la probabilidad, es decir, la frecuencia relativa "teórica" mediante su forma más común de expresión, los porcentajes; así cuando se afirma que "el 20 % de los pacientes con mononucleosis infecciosa presentan el síntoma A", no se hace otra cosa que expresar en lenguaje de los porcentajes un hecho probabilístico relacionado con la práctica médica, en este caso se está expresando en un lenguaje habitual entre los médicos, que la probabilidad de que se presente el síntoma A en un paciente con mononucleosis infecciosa es 0,20 ó lo que es igual, que una medida de la probabilidad de que un paciente con mononucleosis infecciosa presente el síntoma A es de un 0,20.

Su interpretación (la de 0,20) será la de una probabilidad, ya que de hecho este 20 % es el resultado de largas experiencias de observar pacientes con esa enfermedad y de aplicar consecuentemente la ley de regularidad estadística (ver anexo 5).

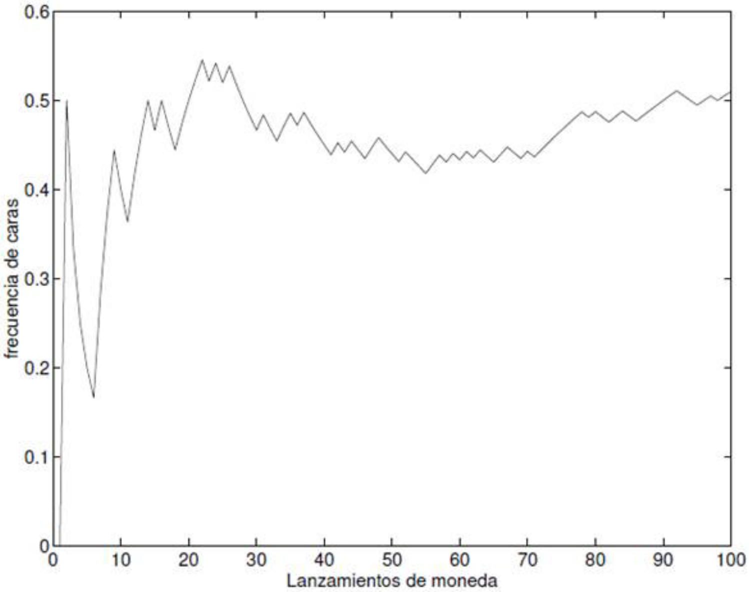
ANEXO 5. Ley de la regularidad estadística

Esta ley fue estudiada en los años nuestros en pregrado pero considero imprescindible traerla para recordar ciertos aspectos básicos.

Retomemos el concepto de fenómeno o experimento aleatorio el cual se denomina aleatorio cuando cumple con la denominada Ley de la regularidad estadística.

La Ley de la regularidad estadística se refiere a determinada estabilidad en la frecuencia relativa observable en los fenómenos aleatorios.

Por ejemplo: observemos en la figura siguiente se presenta la evolución de la frecuencia relativa del número de caras obtenido en el lanzamiento de una moneda en 100 ocasiones (simulado por un ordenador). Primero la evolución de las frecuencias relativas es errática, pero a medida que el número de tiradas aumenta, tiende a lo que entendemos por “probabilidad de cara”.



Pues bien ese gráfico nos da idea clara acerca de la Ley de la regularidad estadística (muestra la tendencia observable en las frecuencias relativas a irse estabilizando alrededor de cierto valor que es su probabilidad, a medida que aumenta el número de observaciones). Esta es una cualidad intrínseca del fenómeno, independiente de quien realice la observación, siempre y cuando permanezcan fijas un conjunto de condiciones.

Entonces a ese valor constante para cada suceso, al cual se acerca cada vez más su frecuencia relativa a medida que aumenta el número de realizaciones del fenómeno o experimento aleatorio es al que se le denomina probabilidad (como se vio en el ejemplo es la “probabilidad de que salga cara”).

Por lo tanto el concepto de probabilidad es un **valor intrínseco de cada fenómeno** y siempre se va a referir a una **gran cantidad de observaciones**; luego se va a poder pronosticar o predecir qué ocurrirá dentro del marco de muchas observaciones. No obstante a partir del valor de probabilidad se puede tener una **idea aproximada** de lo que puede suceder si se toma en cuenta la regularidad estadística.

Cuando el valor de probabilidad asociado a la ocurrencia de un hecho es pequeño se interpreta como que “es improbable que este hecho ocurra”; si por el contrario el valor de probabilidad asociado a la ocurrencia de un hecho es cercana a 1, esto se interpreta como que “este hecho ocurre con gran frecuencia”.

Pero debe quedarles en mente desde ahora que **el concepto de probabilidad se refiere a un valor numérico específico para cada evento** (el que puede estar expresado por medio de una variable aleatoria); en él está implícita la idea de realización de un número grande de observaciones, por lo que ese valor será útil cuando se pretende predecir o pronosticar lo que ocurrirá dentro del ámbito de muchas observaciones y no para resultados aislados. Esto debe quedarles claro a partir de ahora, por eso lo repito. A punto de partida de las probabilidades asociadas a los resultados posibles de una variable aleatoria es imposible predecir exactamente qué ocurrirá en una observación futura. En los fenómenos aleatorios solo se puede tener una idea aproximada de lo que tiene mayores posibilidades de ocurrir.

El valor de probabilidad 0 (cero) es prácticamente imposible de obtener en una experiencia real. Cuando los valores que se obtienen son muy pequeños, para poder decidir si difieren realmente de 0 (cero) o no, se emplean técnicas estadísticas especiales que estudiarán en el módulo Introducción a la Inferencia Estadística.

A manera de resumen:

-Se tiene por un lado que aunque el sentido de la probabilidad está especificado como **el valor teórico** al cual se van aproximando los valores observables de frecuencia relativa de un cierto hecho relacionado con un fenómeno aleatorio y que por tanto puede interpretarse como la “frecuencia relativa teórica” para un número considerablemente grande de realizaciones del fenómeno

-Por otro lado también tiene el sentido de **ser una medida de la posibilidad** de que el hecho asociado con dicho valor de probabilidad pueda o no ocurrir en una sola realización del fenómeno.

Anexo 6. Las variables aleatorias

El origen de la necesidad de aplicar la Bioestadística es la variabilidad de los sucesos biológicos. Esto debe quedarles claro pues es básico para los Bioestadísticos (la existencia de la variabilidad). Si no hubiera variabilidad no se necesitaría la Bioestadística.

Ahora bien, cuando observamos o realizamos un experimento, los datos obtenidos presentan gran variabilidad aún cuando aparentemente esas observaciones o experimentos realizados se realicen en las mismas condiciones.

A diferencia de las propiedades físicas clásicas que se rigen por leyes demostradas, en las ciencias de la salud la repetición de un experimento no lleva necesariamente al mismo resultado.

Por lo tanto en todo proceso de investigación médica el investigador plantea el estudio de variables que no siempre tomarán el mismo valor a pesar de mantener un conjunto de condiciones fijas y estables. Por ejemplo: “el nivel de colesterol en sangre”, si tomamos como conjunto de condiciones fijas en la investigación el que los sujetos bajo estudio sean de una misma edad, masculinos, supuestamente sanos y cubanos, entre otras condiciones, si tomáramos muestras de sangre en sujetos con esas condiciones los resultados sobre el nivel de colesterol no siempre son idénticamente iguales, es decir, presentan ciertas diferencias producto de la influencia de otras condiciones que no son posibles controlar y que se manifiestan en el momento de hacer las diferentes mediciones del nivel de colesterol. Esas son “variables aleatorias” (nivel de colesterol en sangre, por ejemplo) y aquí es necesario que comprendan e interioricen bien que son “variables” porque toman diferentes valores pero lo más importante es que son “aleatorias” porque reflejan el conjunto de factores no controlables que causan las diferencias en las mediciones de esas variables, a pesar de mantener una serie de condiciones fijas.

Esas diferencias en los resultados obtenidos, a pesar de que en unas pocas de ellas hay la apariencia de una total anarquía, de un gran desorden, de no haber sujeción a ningún tipo de ley, cuando estas son analizadas en grandes cantidades por métodos adecuados, se observa que cumplen con determinadas regularidades expresables a través de leyes, las cuales se denominan “Leyes estadísticas o no determinísticas” (ver el anexo 5). Estas leyes se encargan de estudiar la probabilidad en provecho de la Bioestadística.

Estas leyes probabilísticas son de gran importancia pues permiten determinar el comportamiento de esas variables desde el punto de vista de un modelo matemático que posibilitará poder realizar previsiones respecto a su posible comportamiento en las poblaciones de las que procede la muestra en la que se midieron esas variables aleatorias.

Por ejemplo ¿qué valores de esa variable aleatoria en estudio (nivel de colesterol en sangre) se deben emplear para representar a los individuos supuestamente sanos? Más adelante lo veremos.

Por tanto no es posible conocer qué valores tomará una variable aleatoria hasta después de realizada su medición, pero sí se le podrán asignar probabilidades a los distintos valores que puede exhibir.

Entonces ya podemos ir uniendo lo aprendido y deben ya tener claro que el conjunto de todos los posibles resultados de un fenómeno o experimento aleatorio recibe el nombre de población, denominándose elementos a cada uno de sus componentes y pudiendo ser su número finito, o en teoría, infinito. Cabe, también, contemplar una población como un conjunto de elementos con una o más características comunes (el mismo concepto que dieron cuando cursaron el módulo de Estadística descriptiva). La población se caracteriza probabilísticamente mediante variables aleatorias y éstas por sus campos de variación y distribuciones de probabilidad (sobre estas distribuciones se hablará más adelante) que especifican el comportamiento aleatorio de la población. Aquí ven la necesidad de conocer estos aspectos básicos sobre la teoría de las probabilidades.

Anexo 7. Glosario español-inglés de términos empleados en la teoría de las probabilidades

La teoría de las probabilidades es un aspecto de gran importancia dentro de la Bioestadística. Los residentes de Bioestadística cursan este módulo (denominado Probabilidades) durante el primer año de la especialidad. Es necesario que conozcan los conceptos principales de esta materia ya que es la base de esta especialidad.

En ocasiones la literatura disponible sobre probabilidades se encuentra en inglés lo que dificulta su estudio y comprensión si no se tiene un conocimiento de ese lenguaje técnico.

El glosario quedó confeccionado por 45 términos ordenados alfabéticamente; primero se ofrece el término en español y a continuación el término en inglés, como puede verse:

A
Asintóticamente normal / Asymptotic normality

C
Correlación / Correlation Covarianza / Covariance

D
Desviación estándar / Standard deviation
Distribución Bernoulli / Bernoulli distribution
Distribución Binomial / Binomial distribution
Distribución de frecuencias / Frequency distribution
Distribución de probabilidad / Probability distribution
Distribución F de Fisher / F distribution
Distribución muestral / Sampling distribution
Distribución muestral de la media / Sampling distribution of the mean
Distribución normal / Normal distribution
Distribución normal estándar / Standard normal distribution

E
Error cuadrático medio / Mean square error
Error estándar de la media / Standard error of the mean
Espacio muestral / Sample space
Esperanza de una variable aleatoria / Expectation of a random variable
Estimador (estadístico) / Estimator (statistic)
Estimador sesgado / Biased estimator
Evento (suceso) / Outcome
Eventos mutuamente excluyentes / Mutually exclusive events
Eventos independientes / Independent events

F
Función de densidad / Density function
Función de distribución acumulada / Cumulative distribution function

G
Grados de libertad / Degrees of freedom

Histograma / Histogram

I
Independencia / Independence

L
Ley débil de los grandes números / Weak law of large numbers

M
Media / Mean
Muestreo aleatorio / Random sampling

P
Probabilidad / Probability
Probabilidad condicional / Conditional probability
Probabilidad total / Law of total probability

R
Regla de la multiplicación para eventos independientes / Multiplication rule for independent events

S
Sesgo / Bias

T
Teorema central del límite / Central limit theorem
Teorema de Bayes / Bayes' theorem (Bayes’ rule)

V
Valor esperado / Expected value
Variabilidad / Variability
Variable / Variable
Variable aleatoria / Random Variables
Variable aleatoria continua / Continuous random variable
Variable aleatoria discreta / Discrete random variable
Variable independiente / Independent variable
Varianza / Variance

Z
Valor Z / Z score