TEMA *4*

PROBABILIDAD

Propósito: Introducir al estudiante en la teoría de las probabilidades y sus aplicaciones prácticas.

Al finalizar el tema el alumno debe estar en la capacidad de:

- Explicar los conceptos básicos de las probabilidades.
- Aplicar de forma correcta las reglas de la probabilidad.
- Conocer y aplicar las reglas de conteo en el cálculo de las probabilidades.

INTRODUCCIÓN

La probabilidad constituye parte importante de nuestra vida cotidiana. En la toma de decisiones personales y administrativas, nos enfrentamos a la incertidumbre y utilizamos la teoría de la probabilidad. Vivimos en un mundo incapaz de predecir el futuro con total certidumbre. Nuestra necesidad de encarar a la incertidumbre nos lleva a estudiar y utilizar la teoría de la probabilidad.

En muchos casos tendremos algún conocimiento sobre los posibles resultados de una decisión. Al organizar esta información y considerarla de manera sistemática seremos capaces de reconocer nuestras suposiciones, comunicar nuestro razonamiento a otras personas y tomar una decisión más sólida que la que tomaríamos si sólo consideramos nuestra intuición.

CONTENIDO

- Introducción a la teoría de la probabilidad. Comparación entre probabilidad y estadística. Nociones básicas. Fenómenos aleatorios vs. fenómenos determinísticos. Términos básicos, reglas y propiedades; eventos mutuamente excluyentes, eventos exhaustivos, eventos independientes. Enfoques para asignar probabilidades: enfoque clásico, enfoque de la frecuencia relativa y enfoque subjetivo.
- Reglas de la probabilidad: Representación de probabilidades usando diagramas de Venn, diagramas de árbol y tablas de contingencia. Probabilidad marginal, probabilidad conjunta y probabilidad condicional. Regla de la adición y regla de la multiplicación. Teorema de Bayes.
- Reglas de conteo: Muestras con repetición y sin repetición. Muestras ordenadas y no ordenadas. Variaciones, Permutaciones, Combinaciones.

DESARROLLO

- Introducción a la teoría de la probabilidad.

En el caso de un hecho considerado, del cual no se puede dar un resultado exacto (en forma determinística) y que por el contrario lo que se tiene es incertidumbre sobre lo observado, nos lleva a pensar que estamos en presencia del concepto de probabilidad.

Es el caso típico de analizar poblaciones por medio de muestras, solo tendríamos la información correcta si hacemos la medición en todos lo elementos que la integran. Como esto en muchos casos no es posible, se efectúa el análisis a partir de muestras, pero de hecho el valor obtenido de la muestra, tendrá una diferencia con el que se obtendría de la población. La relación entre estos valores se establece con la aplicación de la teoría de probabilidad.

- Comparación entre probabilidad y estadística.

Como ya se ha dicho, la estadística consiste de un conjunto de métodos y técnicas para la recolección, organización, representación e interpretación de datos. Por medio de la estadística podemos describir una población a través de ciertos valores medidos o calculados directamente de sus elementos cuando esto es posible, pero en muchos casos el acceso a la población total está restringido o es imposible y se recurre al uso de muestras.

En estadística, para que una muestra sea de utilidad debe ser representativa de la población y esto se consigue por medio del muestreo aleatorio, es decir cuando no se manipula la muestra a conveniencia, sino que su obtención es completamente al azar, por supuesto considerando que todos los elementos de la población tienen la misma probabilidad de ser elegidos.

En consecuencia, hemos de notar que la probabilidad es un concepto importante para la estadística. La probabilidad es definida por muchos autores como la posibilidad de ocurrencia de un suceso o evento, pero su cálculo requiere el estudio toda una teoría. Aquí en estadística, elegir un elemento de la población es el suceso del cual se desea que tenga la misma probabilidad de ser elegido a los demás. Queda entonces comprender como se calcula la probabilidad y sus nociones básicas.

- <u>Nociones básicas.</u>

Definición de probabilidad

En términos generales, la probabilidad es un número entre cero y uno, inclusive, que describe la posibilidad relativa (oportunidad o casualidad) de que ocurra un evento.

Es común que la probabilidad sea expresada en forma decimal, como 0.70, 0.27 o 0,50. No obstante, también se da en forma de fracción, como 7/10, 27/100 o ½, y también en forma de porcentaje, 70%, 27% o 50%.

Cuanto más próxima se encuentre una probabilidad a 0, más improbable es que el evento suceda. Cuanto más próxima se encuentre la probabilidad a 1, más seguro es que suceda.

Tres palabras clave: experimento, resultado y evento

Estos términos son empleados en el lenguaje de la vida cotidiana, pero en estadística adquieren significados específicos.

EXPERIMENTO: Proceso que induce a que ocurra una y sólo una de varias posibles observaciones.
RESULTADO: Observación o respuesta particular de un experimento.
EVENTO: Conjunto de uno o más resultados de un experimento.

Por ejemplo:

Experimento:	Lanzamiento de un dado
Resultados:	1, 2, 3, 4, 5 y 6
Eventos:	Se observa un número par; Se observa un número mayor que 4; Se observa el número 3.

- Fenómenos aleatorios vs. fenómenos determinísticos.

Los ***fenómenos deterministas*** al repetirlo en idénticas condiciones se obtiene el mismo resultado, mientras que el ***fenómeno aleatorio*** no es posible predecir su resultado. La estadística se ocupa de aquellos fenómenos no deterministas donde es imposible predecir los resultados.

- Términos básicos, reglas y propiedades
  - <u>Eventos mutuamente excluyentes</u>

MUTUAMENTE EXCLUYENTE El hecho de que un evento se presente significa que ninguno de los demás eventos puede ocurrir al mismo tiempo.

  - <u>Eventos exhaustivos</u>

COLECTIVAMENTE EXHAUSTIVO Por lo menos uno de los eventos debe ocurrir cuando se lleva a cabo un experimento.

- Eventos independientes

Si un evento ocurre, no tiene ningún efecto sobre la probabilidad de que otro evento acontezca.

- Enfoques para asignar probabilidades:

Conviene analizar dos perspectivas para asignar probabilidades: los enfoques objetivo y subjetivo. La probabilidad objetiva se subdivide en a) probabilidad clásica y b) probabilidad empírica.

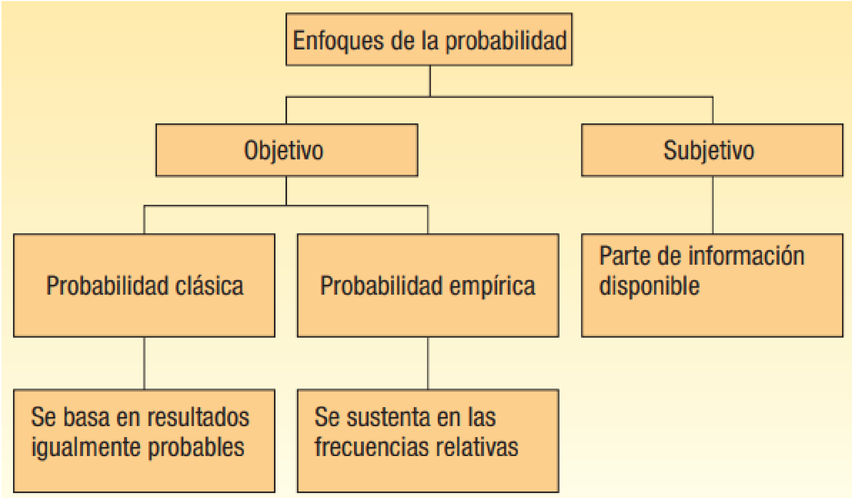


Figura: Resumen de enfoques de la probabilidad

- Enfoque clásico

$$\text{Probabilidad de un evento} = \frac{\text{Número de resultados favorables}}{\text{Número total de posibles resultados}}$$

- Enfoque de la frecuencia relativa

La probabilidad empírica o frecuencia relativa, el segundo tipo de probabilidad, se basa en el número de veces que ocurre el evento como proporción del número de intentos conocidos.

PROBABILIDAD EMPÍRICA La probabilidad de que un evento ocurra representa una fracción de los eventos similares que sucedieron en el pasado.

$$\text{Probabilidad empírica} = \frac{\text{Número de veces que el evento ocurre}}{\text{Número total de observaciones}}$$

- Enfoque subjetivo

CONCEPTO SUBJETIVO DE PROBABILIDAD Posibilidad (probabilidad) de un evento en particular que asigna un individuo a partir de cualquier información que encuentre disponible.

Algunos ejemplos de probabilidad subjetiva son los siguientes:

1. Calcular la posibilidad de que los Patriotas de Nueva Inglaterra jueguen el Súper Tazón el año que viene.
2. Calcular la posibilidad de que usted contraiga matrimonio antes de los 30 años.
3. Calcular la posibilidad de que el déficit presupuestario de Venezuela se reduzca a la mitad en los siguientes 10 años.

- <u>Reglas de la probabilidad:</u>
    - Probabilidad marginal, probabilidad conjunta y probabilidad condicional.

Probabilidad marginal. Es la probabilidad de que ocurra un solo evento.

Probabilidad conjunta. Es cuando dos eventos ocurren al mismo tiempo.

El siguiente diagrama de Venn muestra dos eventos que no son mutuamente excluyentes. Ambos se superponen para ilustrar el evento conjunto de que algunas personas hayan visitado ambas atracciones.

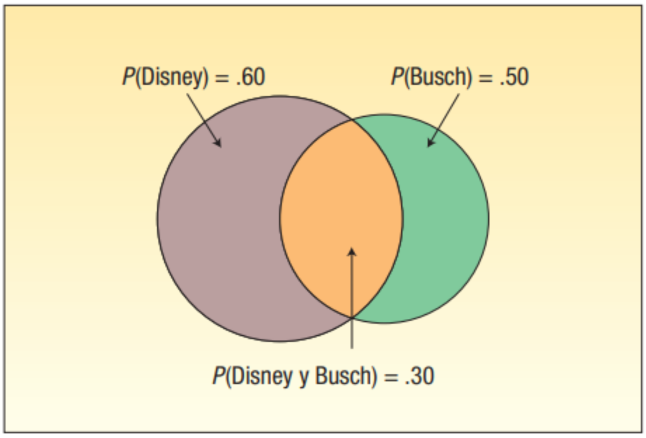


Figura. Eventos no mutuamente excluyentes, pueden ocurrir al mismo tiempo.

Así, la *Probabilidad conjunta* que mide la posibilidad de que dos o más eventos sucedan simultáneamente.

Probabilidad condicional. Probabilidad de que un evento en particular ocurra, dado que otro evento haya acontecido. Cuando dos eventos no son independientes, se dice que son ***dependientes***. En este caso, los eventos *son dependientes*, ya que la ocurrencia de uno cambia la probabilidad de ocurrencia del otro.

Para representar la probabilidad condicional del evento A dado que ocurrió el evento B, se escribe: P(A|B).

- <u>Regla de la adición y regla de la multiplicación.</u>

Regla especial de la adición. Si dos eventos A y B son mutuamente excluyentes, la regla especial de la adición establece que la probabilidad de que ocurra uno u otro es igual a la suma de sus probabilidades.

REGLA ESPECIAL DE LA ADICIÓN $P(A \text{ o } B) = P(A) + P(B)$

Regla del complemento. Se emplea para determinar la probabilidad de que un evento A ocurra restando de 1 la probabilidad del evento complemento de A (que se escribe ~A) el cual no ha ocurrido.

REGLA DEL COMPLEMENTO $P(A) = 1 - P(\sim A)$

Regla general de la adición. Los resultados de un experimento pueden no ser mutuamente excluyentes. En este caso, la regla para dos eventos designados A y B se escribe así:

REGLA GENERAL DE LA ADICIÓN $P(A \text{ o } B) = P(A) + P(B) - P(A \text{ y } B)$

Regla de la multiplicación. Se utiliza para determinar la probabilidad de dos eventos que se presentan simultáneamente. Al igual que para la adición, se tienen dos reglas: la regla especial y la regla general:

Regla especial de la multiplicación. Requiere que dos eventos, A y B, sean ***independientes***, lo cual quiere decir que la ocurrencia de un evento no tiene ningún efecto sobre la probabilidad de ocurrencia del otro. En este caso, la probabilidad de que A y B ocurran se determina multiplicando las dos probabilidades.

REGLA ESPECIAL DE LA MULTIPLICACIÓN $P(A \text{ y } B) = P(A)P(B)$

Regla general de la multiplicación. Esta regla sirve para determinar la probabilidad conjunta de dos eventos cuando éstos ***no son independientes***. Simbólicamente, la probabilidad conjunta, P(A y B), se calcula de la siguiente manera:

REGLA GENERAL DE LA MULTIPLICACIÓN $P(A \text{ y } B) = P(A)P(B|A)$

La regla general de la multiplicación establece que, en caso de dos eventos, A y B, la probabilidad conjunta de que ambos eventos ocurran se determina multiplicando la probabilidad de que ocurra el evento A por la probabilidad condicional de que ocurra el evento B, dado que A ha ocurrido.

- Representación de probabilidades usando
  - Tablas de contingencia, diagramas de árbol y Diagramas de Venn.

Tabla de contingencia.

Como ya se sabe, la tabla de contingencia es una tabulación cruzada, que resume simultáneamente dos variables de interés. Se utiliza para clasificar observaciones (de escala nominal u ordinal) de acuerdo con dos características identificables que representan variables cualitativas o categóricas. Con el uso de tablas de contingencia se facilita determinar las probabilidades marginales, conjuntas, totales o especificas de las variables.

En la siguiente figura se muestra una ***tabla de contingencia*** con información sobre la *lealtad de los ejecutivos* y el *tiempo de servicio a una compañía*, que se tomará como ejemplo en el próximo punto para calcular distintas probabilidades.

	Tiempo de servicio				
Lealtad	Menos de 1 año, B_1	1 a 5 años, B_2	6 a 10 años, B_3	Más de 10 años, B_4	Total
Permanecería, A_1	10	30	5	75	120
No permanecería, A_2	25	15	10	30	80
	35	45	15	105	200

Figura. Tabla de contingencia de los tiempos de servicio y la lealtad de los ejecutivos a la compañía.

Diagrama de árbol. Es una gráfica útil para organizar cálculos que implican varias etapas. Cada segmento del árbol constituye una etapa del problema. Las ramas del árbol se ponderan por medio de probabilidades. Con frecuencia, se vale de una tabla de contingencia donde se tiene la información resumida.

El diagrama de árbol es útil entonces para responder preguntas como: ¿Cuál es la probabilidad de seleccionar al azar a un ejecutivo leal a la compañía —que permanecería en ella— y cuál de ellos tiene más de 10 años de servicio?

Para explicar su construcción, se hará referencia a la tabla contingencia presentada en el punto anterior, sobre los *tiempos de servicio y la lealtad de los ejecutivos a una compañía*. Entonces, los pasos a seguir para construir un diagrama de árbol son:

1. Comenzar dibujando un punto grueso a la izquierda para representar la raíz del árbol (vea la figura del diagrama de árbol de la página siguiente).

2. En este problema, dos ramas principales salen de la raíz: la rama superior representa el evento "permanecería" y la rama inferior el evento "no permanecería". Sus probabilidades se anotan sobre las ramas, en este caso, 120/200 y 80/200. Estas probabilidades también se denotan P(A1) y P(A2).

3. De cada una de las ramas principales salen cuatro ramas, las cuales representan el tiempo de servicio: menos de 1 año, 1 a 5 años, 6 a 10 años y más de 10 años. Las probabilidades condicionales de la rama superior del árbol, 10/120, 30/120, 5/120, etc., se anotan en las ramas adecuadas, que son P(B1|A1), P(B2|A1), P(B3|A1) y P(B4|A1), en las cuales B1 se refiere a menos de 1 año de servicio; B2, a 1 a 5 años de servicio, B3, a 6 a 10 años de servicio y B4, a más de 10 años. En seguida, anotamos las probabilidades condicionales en la rama inferior.

4. Por último, las probabilidades conjuntas relativas al hecho de que los eventos A1 y Bi o los eventos A2 y Bi ocurrirán al mismo tiempo aparecen al lado derecho. Por ejemplo, de acuerdo con la fórmula (5-6), la probabilidad conjunta de seleccionar al azar a un ejecutivo que permanecería en la compañía y que tenga más de 1 año de servicio es:

$$P(A_1 \text{ y } B_1) = P(A_1)P(B_1|A_1) = \left(\frac{120}{200}\right)\left(\frac{10}{120}\right) = 0.05$$

Como las probabilidades conjuntas representan todos los posibles resultados (permanecería, 6 a 10 años de servicio, no permanecería, más de 10 años de servicio, etc.), deben sumar 1.00. (vea la gráfica siguiente).

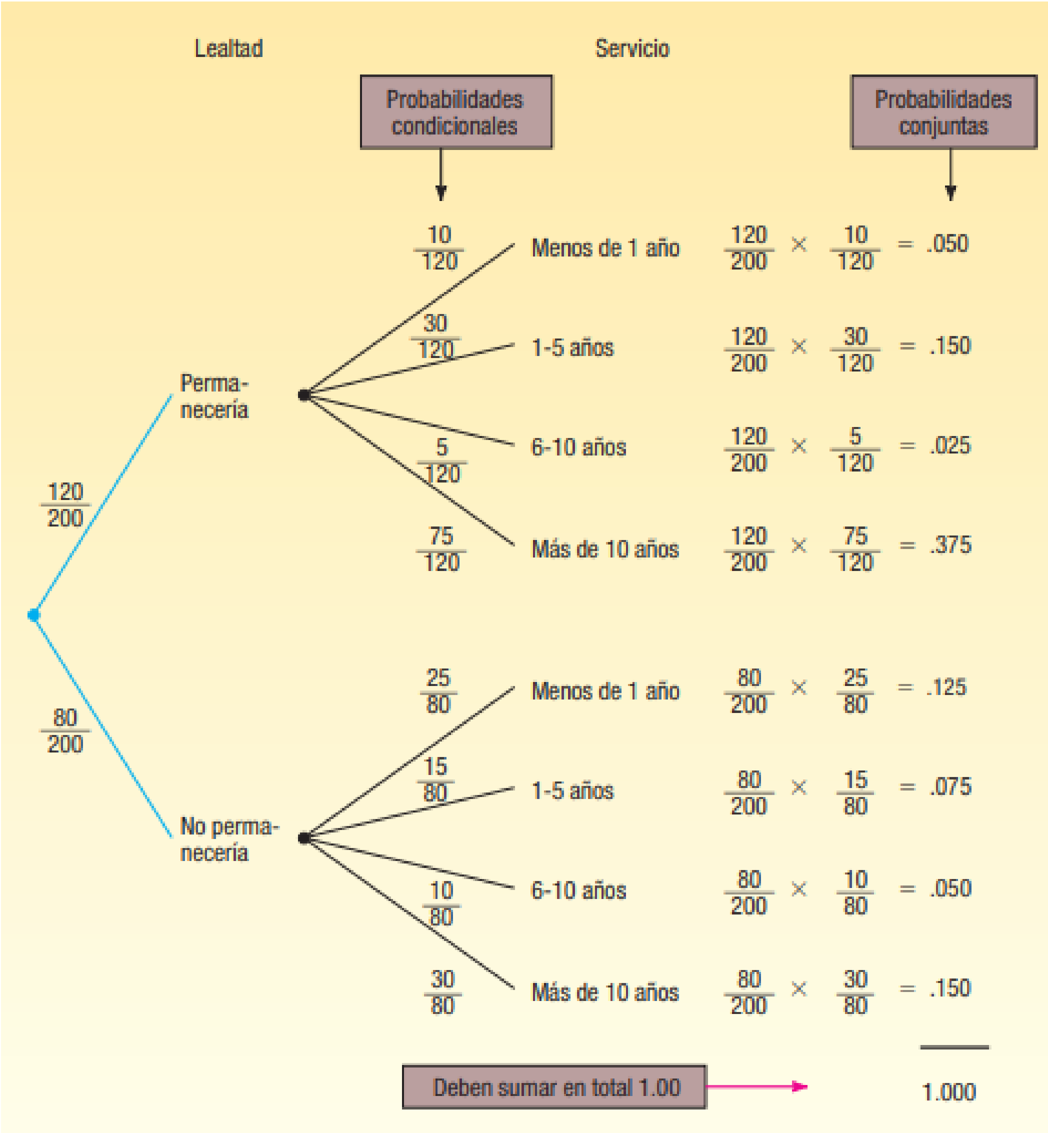


Figura. Diagrama de árbol que muestra la lealtad y los años de servicio.

El ***diagrama de Venn*** es otro de los diagramas utilizados con frecuencia en la Estadística y las probabilidades. Fue credo para representar de manera gráfica los resultados de un experimento.

Para construir un diagrama de Venn, primero se encierra un espacio, el cual representa el total de posibles resultados. Este espacio es de forma rectangular. Así, un evento se representa por medio de un área circular, que se dibuja dentro del rectángulo, la cual corresponde a la probabilidad del evento.

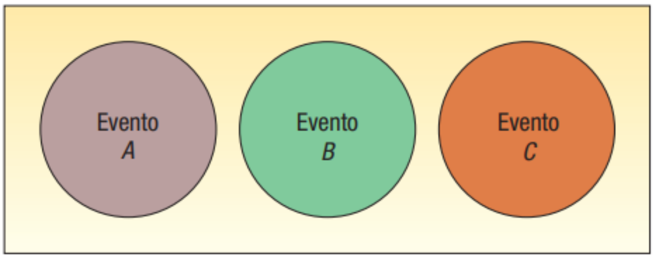
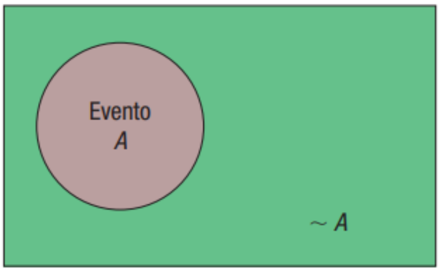
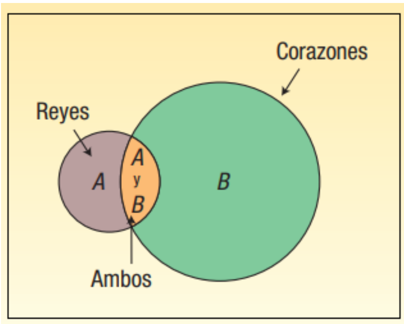


Figura. Diagrama de Venn con tres eventos igualmente probables.

Por ejemplo, la regla del complemento puede ser representada mediante el siguiente diagrama de Venn:



En este otro caso, el diagrama de Venn muestra dos eventos que no son mutuamente excluyentes. Ambos se superponen para ilustrar el evento conjunto de que algunas personas hayan visitado ambas atracciones.



- Teorema de Bayes.

El teorema de Bayes es un método que consiste en revisar una probabilidad, dado que se ha logrado información adicional. En el caso de dos eventos mutuamente excluyentes y colectivamente exhaustivos, se utiliza la siguiente fórmula:

$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)}$$

Ejemplo:

En la Uneg, la probabilidad de que a un alumno seleccionado al azar le guste las **clases presenciales** es del 60 %, mientras que la probabilidad de que a un alumno le guste las **clases online** es del 36 %. Además, se sabe que la probabilidad de que a un alumno le guste las **clases online dado** que le gusta las **clases presenciales** es del 40%. Calcular la probabilidad de que a un alumno le guste las clases presenciales, dado que le gusta las clases online.

Para dar respuesta a este problema, primero definimos los dos eventos con los que vamos a trabajar:

p: que a un alumno le guste las clases presenciales.
o: que a un alumno le guste las clases online.

Tenemos los siguientes datos: $P(p) = 0{,}6$ $P(o) = 0{,}36$ $P(o|p) = 0{,}4$

Nos piden calcular: $P(p|o)$.

Aplicamos el teorema de Bayes:

$$P(\mathrm{p}|\mathrm{o}) = \frac{P(\mathrm{p}) \cdot P(\mathrm{o}|\mathrm{p})}{P(\mathrm{o})}$$

$$P(\mathrm{p}|\mathrm{o}) = \frac{0{,}6 \cdot 0{,}4}{0{,}36} = \frac{0{,}24}{0{,}36} = \frac{\cancel{24}}{\cancel{36}} = \frac{\mathbf{2}}{\mathbf{3}} = \mathbf{0{,}6667} = \mathbf{66{,}67\ \%}$$

Entonces, la probabilidad de que un alumno(a) le guste las clases presenciales dado que le gusta las clases online es de 0,6667 o 66,67 %

- Reglas de conteo:
    - Variaciones, Permutaciones, Combinaciones.

Existen dos tipos de variaciones: con repetición y sin repetición.

Variaciones con repetición: Cuando dentro de cada tupla se puede repetir un elemento más de una vez.

Las variaciones con repetición se calculan mediante la fórmula: N^n

(donde N es el tamaño de la población y **n** el número de elementos de la muestra).

Variación sin repetición: Significa que los elementos no se pueden repetir dentro de la misma tupla.

En este caso, la fórmula a utilizar es: N!/(N-n)!

(donde N es el tamaño de la población y **n** el número de elementos de la muestra).

Existen tres reglas de conteo útiles para determinar el número de resultados de un experimento.

A. **La regla de la multiplicación** establece que si hay ***m*** formas de que un evento suceda y ***n*** formas de que otro pueda suceder, entonces hay ***mn*** formas en que los dos eventos pueden suceder.

Número de disposiciones = (m)(n)

B. Una **permutación** es un arreglo en el que el orden de los objetos seleccionados de un conjunto específico es importante.

$${}_nP_r = \frac{n!}{(n-r)!}$$

C. Una **combinación** es un arreglo en el que el orden de los objetos seleccionados de un conjunto específico no es importante.

$${}_nC_r = \frac{n!}{r!(n-r)!}$$

- Muestras con repetición y sin repetición.

Una muestra puede ser considerada como un subconjunto de grupos o tupla de **n** elementos {e1, e2, …, en} extraídos de una población.

Las **muestras con repetición** (o reposición) permiten que cualquier elemento de la población pueda ser observado más de una vez. En este caso el tamaño **N** de la población no cambia.

Cuando las **muestras** son **sin reposición,** el elemento observado se saca de la población y no es considerado de nuevo. La población en este caso se reduce, en otras palabras, cambia de tamaño.

- Muestras ordenadas y no ordenadas.

Si las **muestras son ordenadas**, las tuplas que la forman debe considerar la posición de los elementos. En este caso, elementos como (e_n, e_m) se considera diferente a (e_m, e_n). Por ejemplo, (1,2) es diferente a (2,1). Por lo tanto, cada uno cuenta y es diferentes al otro.

Cuando las **muestras no son ordenadas**, no importa el orden en que fueron extraídas de la población. Por ejemplo, considere una muestra conformada por tres números: 2, 4 y 5. En este caso, cualquiera de los siguientes resultados se considera idéntico: (2, 4, 5), (2, 5, 4), (4, 2, 5), (4, 5, 2), (5, 2, 4), (5, 4, 2) y debe contarse como una sola muestra.