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 es un concepto importante en nuestras vidas; sobre todo, cuando se deben tomar decisiones importantes a nivel personal, en ambientes laborales y en muchas otras situaciones en las que existe razón para dudar del resultado obtenido por causa de nuestra elección, simplemente, por factores aleatorios o que no están bajo nuestro control. En general, vivimos en un mundo incapaz de predecir el futuro con total seguridad.

La necesidad de encarar a la incertidumbre nos lleva a estudiar la aleatoriedad. La teoría de la probabilidad se encarga de esto y nos ofrece métodos y procedimientos confiables que pueden ser utilizados para anticipar los resultados, medir el riesgo y obtener así una ventaja significativa.

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 los 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.

En muchos casos se conocerán los posibles resultados que se pueden obtener cada vez que se realiza una cierta actividad o procedimiento. Sin embargo, habrá casos en los que realizar dicho procedimiento no garantiza que se obtendrá el resultado particular que esperamos entre todos los posibles que podrían ocurrir. Esta situación, un tanto desconcertante, es lo que se denomina un proceso estocástico o no determinista. Por medio de las probabilidades se analizan tales situaciones y se desarrollan métodos matemáticos que permitirán demostrar sistemáticamente algunos patrones regulares que pueden anticipar el comportamiento o incidencia de posibles resultados esperados que caracterizan el conjunto del cual provienen.

De este modo, la teoría de las probabilidades se convierte en una herramienta poderosa para la toma de decisiones, pues a partir de un proceso estocástico (proceso, fenómeno, procedimiento o experimento aleatorio) seremos capaces de evaluar posibles conjeturas, comunicar nuestro razonamiento a otras personas y señalar la respuesta idónea (de mayor probabilidad y menor riesgo) demostrable matemáticamente, aunque subjetivamente no sea la que indica nuestra intuición.

Por otra parte, 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, tomar una muestra que sea representativa de la población implica que no sea manipulada a conveniencia, sino que su obtención sea completamente al azar. Este método, conocido como muestreo aleatorio, considera que todos los elementos de la población tienen la misma oportunidad de ser elegidos. En tal sentido, es un proceso estocástico (un experimento aleatorio, no determinístico), el cual puede ser analizado por la teoría de las probabilidades. De este modo, la estadística y las probabilidades quedan unidas para el análisis e interpretación de fenómenos aleatorio, no predecibles en lo individual, pero si a nivel de un colectivo (ley de los grandes números), lo cual la convierte en una herramienta útil para la toma de decisiones acertadas sobre una base matemática demostrable.

Nociones básicas.

Definición de probabilidad

En términos generales, la probabilidad es un número real que pertenece al conjunto [0, 1] y describe la posibilidad relativa (respecto a los resultados posibles) de que ocurra un evento.

Es común que la probabilidad sea expresada en forma decimal (p.ej. 0.70, 0.27 o 0,50). No obstante, también se da en forma de fracción (p. ej. 7/10, 27/100 o ½), o bien, en forma porcentual (p.ej. 70%, 27% o 50%).

- Cuanto más próxima se encuentre una probabilidad a 0, más difícil 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 coloquial, pero en estadística adquieren significados específicos.

- ***Experimento***: Proceso que induce a que ocurra una y sólo una de varias posibles observaciones.
  Por ejemplo: Lanzamiento de un dado

- ***Resultado***: Observación o respuesta particular de un experimento.
  Por ejemplo: 1, 2, 3, 4, 5 y 6

- ***Evento*** *(o Suceso)*: Conjunto de uno o más resultados de un experimento.
  Por ejemplo: Se observa un número par; Se observa un número mayor que 4; Se observa el número 3.

Un *Experimento aleatorio*, es cualquier procedimiento (o fenómeno) del cual no se puede anticipar el resultado, aunque su realización se repita en igualdad de condiciones. Esto, debido a que el resultado depende del azar.

Fenómenos aleatorios vs. fenómenos determinísticos.

De forma muy general, cualquier situación o fenómeno observable puede ser clasificado en dos tipos: determinístico y no determinístico (o aleatorio).

- ***Fenómenos deterministas*** son aquellos que al repetirlo en idénticas condiciones se obtiene siempre el mismo resultado.
- ***Fenómeno aleatorio*** son aquellos en los que no es posible predecir su resultado con absoluta certeza.

Las Probabilidad se ocupa de los fenómenos aleatorios (no deterministas) donde es imposible predecir los resultados.

Términos básicos, reglas y propiedades

Eventos mutuamente excluyentes, eventos exhaustivos, eventos independientes

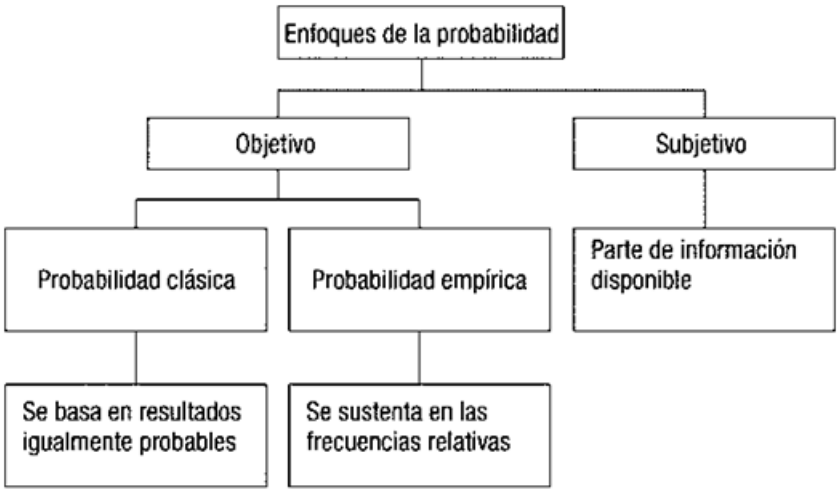
- Eventos ***Mutuamente Excluyentes***: El hecho de que un evento se presente significa que ninguno de los demás eventos puede ocurrir al mismo tiempo.

- Los eventos son ***Colectivamente Exhaustivos*** cuando la suma de las probabilidades de todos los eventos posibles es igual a 1, lo que indica que se han considerado todas las posibilidades.

- ***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.



<u>Enfoque clásico</u>

La definición clásica de probabilidad es la razón entre el número de ocurrencias de un evento (resultados favorables) y el número total de casos posibles (total de posibles resultados).

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

<u>Enfoque de la frecuencia relativa</u>

La probabilidad de que un evento ocurra representa una fracción de los eventos similares que sucedieron en el pasado. Se calcula dividiendo el número de veces que el evento curre entre el total de observaciones.

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

<u>Enfoque subjetivo</u>

Se define como la posibilidad (probabilidad) de un evento en particular ue 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 estudiantes de la UNEG se gradúen Cum Laude.
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.

Reglas de la probabilidad

Representación de probabilidades usando diagramas de Venn, diagramas de árbol y tablas de contingencia.

- ***Los diagramas de Venn*** son círculos que se superponen para mostrar las relaciones lógicas entre dos o más conjuntos de elementos. En Probabilidades, el diagrama de Venn se utiliza frecuentemente para representar eventos. Por ejemplo, la siguiente figura muestra dos eventos que no son mutuamente excluyentes. Se trata del evento de visitar Disney y el evento de visitar el parque Busch. Puede verse en esta figura que ambos eventos se superponen para ilustrar que se trata de un evento conjunto en el que algunas personas han visitado las dos atracciones turísticas que se indican allí.

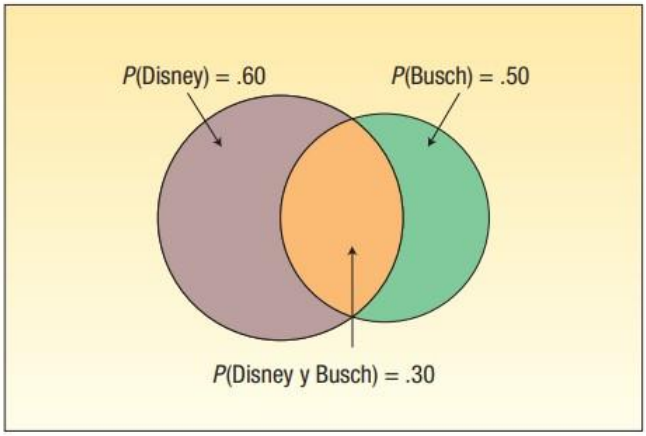
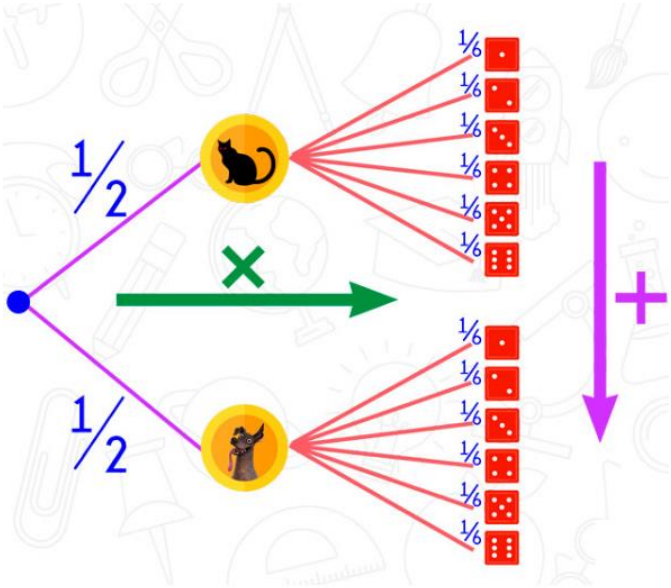


Figura. Eventos no mutuamente excluyentes, pueden ocurrir al mismo tiempo.

- Un ***diagrama de árbol*** es un gráfico que se utiliza para representar los resultados de un experimento aleatorio y determinar las probabilidades de cada resultado. Las ramas del diagrama se identifican con frecuencias o probabilidades, y para calcular las probabilidades se multiplican las probabilidades de ramas adyacentes y se suman las probabilidades de ramas separadas que emergen de un mismo punto.



- Las ***tablas de contingencia*** consisten en un arreglo bidimensional con varias filas y columnas representando los valores posibles de dos variables categóricas relacionadas e indicando la frecuencia con que se repiten tanto de forma independiente como en la forma conjunta o condicional.

CONTRIBUYENTE	1ER TRIMESTRE	2DO TRIMESTRE	3ER TRIMESTRE	4TO TRIMESTRE	TOTALES
Con declaración	A1	A2	A3	A4	**∑Ai**
Sin declaración	B1	B2	B3	B4	**∑Bi**
TOTALES	**A1+B1**	**A2+B2**	**A3+B3**	**A4+B4**	**∑(Ai+bi)**

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.
- **Probabilidad condicional**. Es cuando un evento ocurre, dado que otro evento haya acontecido.

La probabilidad condicional establece la dependencia entre los eventos. Así, cuando dos eventos son dependientes, la ocurrencia de uno cambia la probabilidad de ocurrencia del otro. De otro modo, se dice que los eventos son independientes.

Para representar la probabilidad condicional del evento A dado que ocurrió el evento B, se escribe P(A|B) y su cálculo se obtiene mediante el cociente entre la probabilidad conjunta y la probabilidad marginal del evento impuesto como condición.

Las tablas de contingencias son especialmente útiles para calcular estas probabilidades. Así, en la tabla de contingencia, las *probabilidades marginales* se consiguen al dividir los totales de cada columna o fila entre el total general, las *probabilidades conjuntas* al dividir cada valor entre el total general y las *probabilidades condicionales* al dividir cada valor entre el total de su fila o columna.

Regla de la adición y regla de la multiplicación.

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 $$P(A \text{ y } B) = P(A)P(B)$$

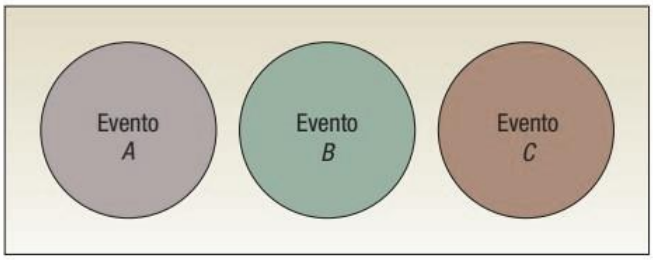
Regla general de la multiplicación. Esta regla sirve para determinar la probabilidad conjunta de dos eventos cuando éstos ***no son independientes***. Simbólicamente, la probabilidad conjunta, P(A y B), se calcula de la siguiente manera:

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

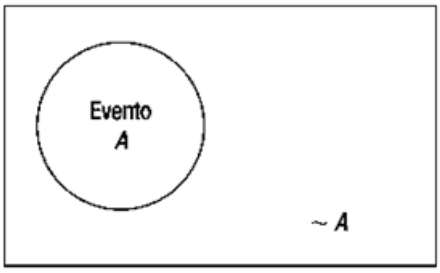
La regla general de la multiplicación establece que, en caso de dos eventos, A y B, la probabilidad conjunta de que ambos eventos ocurran se determina multiplicando la probabilidad de que ocurra el evento A por la probabilidad condicional de que ocurra el evento B, dado que A ha ocurrido.

<u>Cálculo de probabilidades utilizando Diagramas de Venn, Diagramas de árbol y Tablas de contingencia</u>

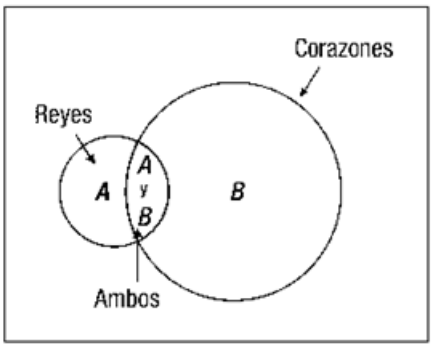
El ***diagrama de Venn*** es uno de los diagramas utilizados con frecuencia en la Estadística y las probabilidades. Fue credo para representar de manera gráfica los resultados de un experimento. Para construir un diagrama de Venn, primero se encierra un espacio, el cual representa el total de posibles 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.



Por ejemplo, la regla del complemento puede ser representada mediante el siguiente diagrama de Venn. Esta regla indica que el suceso contrario o complemento de A es aquel que incluye todos los sucesos elementales que no están en A. $P(A) + P(\bar{A}) = 1$



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.



Una ***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 el cálculo de las probabilidades marginales, conjuntas, totales o específicas 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**
Permaneceria, A_1	10	30	5	75	120
No permaneceria, 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.

Un ***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?

El siguiente ejemplo explica cómo se construye el diagrama de árbol tomando como referencia la tabla de contingencia presentada en el punto anterior:

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).

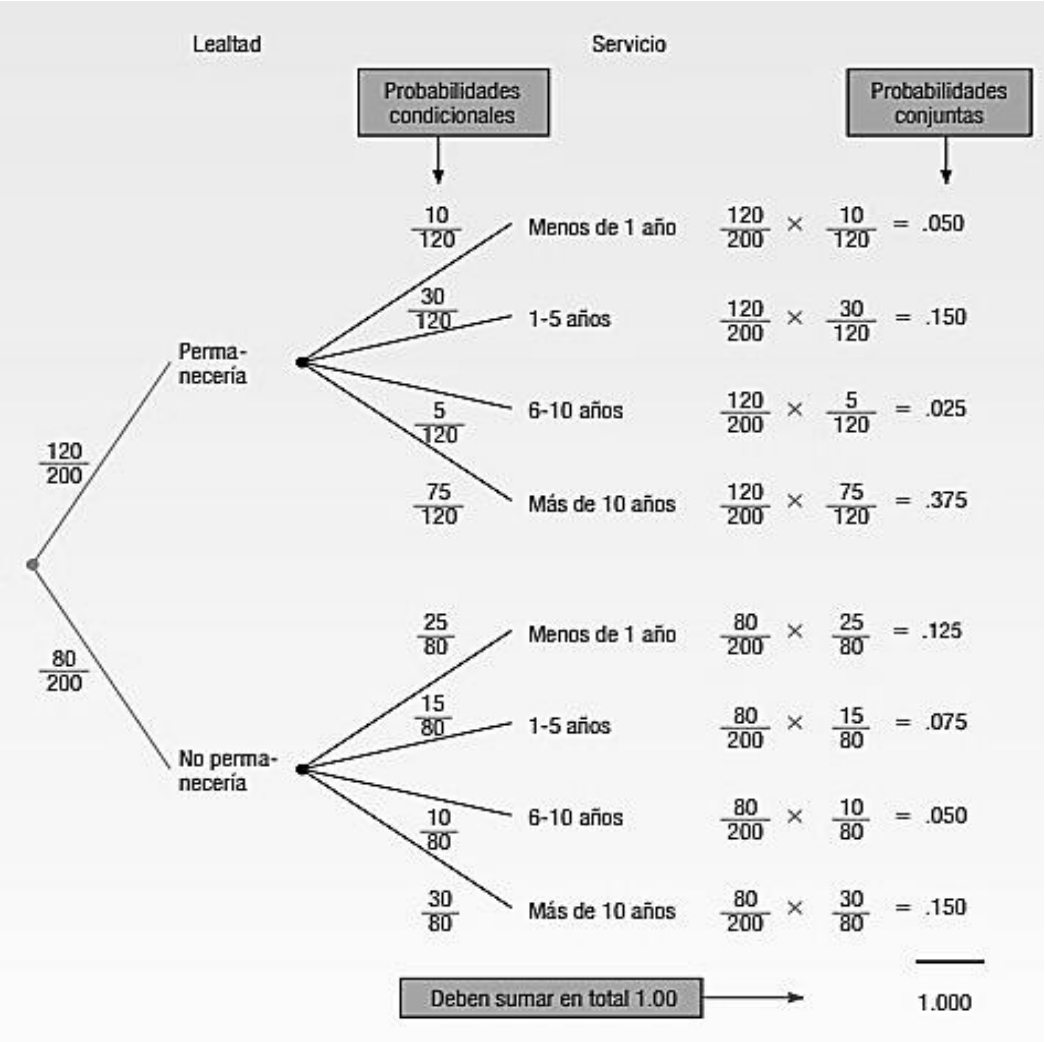


Figura. Diagrama de árbol que muestra la lealtad y los años de servicio.

Teorema de Bayes.

El teorema de Bayes es un método que consiste en revisar una probabilidad, dado que se ha logrado información adicional. Expresa la probabilidad de que ocurra A, dado que ha ocurrido B, en función de la probabilidad de B, dado que ha ocurrido A, por la probabilidad de A, entre la probabilidad de B.

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

En caso que el suceso A represente dos sucesos, A1 y A2, mutuamente excluyentes y colectivamente exhaustivos, entonces la fórmula de Bayes para hallar $P(A_1|B)$ será:

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

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 virtuales es del 36%. Además, se sabe que la probabilidad de que a un alumno le guste las clases virtuales 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 virtuales.

Solución:

Para dar respuesta a este problema, primero se definen los eventos:

A: al alumno le gustan las clases presenciales.
B: al alumno le gustan las clases virtuales.

Entonces, se tiene como datos:

$P(A) = 0{,}60$ $P(\bar{A}) = 0{,}40$
$P(B) = 0{,}36$
$P(B|A) = 0{,}40$

Se pide calcular: ¿$P(A|B)$ =?

Aplicamos el teorema de Bayes: $P(A|B) = P(A)\,P(B|A) / P(B) = = 0{,}60 \times 0{,}40 / 0{,}36 = 0{,}6667$

Interpretación:

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

Reglas de conteo

Muestras con repetición y sin repetición.

Una muestra puede ser considerada como un subconjunto de grupos o tupla de **n** elementos $\{e_1, e_2, \ldots, e_n\}$ 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, la pareja (1, 2) es diferente a (2, 1). Por lo tanto, cada par se cuenta por separado y se consideran diferentes.

En cambio, 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.

Variaciones, Permutaciones, Combinaciones.

<u>Existen dos tipos de variaciones: con repetición y sin repetición.</u>

Variaciones con repetición: Es cuando se puede repetir un elemento más de una vez dentro de cada tupla. 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)!

<u>Existen tres reglas de conteo útiles para determinar el número de resultados de un experimento.</u>

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