PROBABILIDAD

UNIDAD 1. FUNDAMENTOS DE PROBABILIDAD

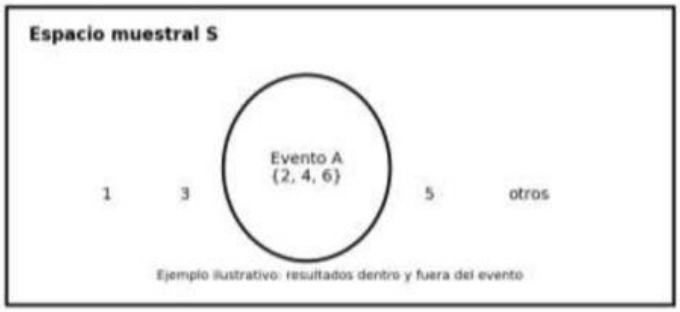
1.1 Espacio muestral y eventos

Experimento aleatorio: es una acción cuyo resultado no puede conocerse con certeza antes de realizarla, aunque se conocen los resultados posibles.

Ejemplos: elegir un producto al azar, seleccionar un trabajador o lanzar un dado.

Espacio muestral (S): es el conjunto de todos los resultados posibles. Un evento (A) es uno o varios resultados que cumplen una condición.

$$S = \{\text{todos los resultados posibles}\} \quad A \subseteq S$$



- Cuando se necesita enumerar los resultados posibles de una selección o experimento.
- Cuando se define una condición: elegir un producto azul, un trabajador técnico o una aplicación con error.
- Es el primer paso antes de calcular probabilidades.

Fórmula: Regla de Laplace:

$$P(A) = \frac{\text{Casos favorables}}{\text{Casos posibles}}$$

Datos:

- Casos favorables: representa los resultados favorables
- Casos posibles: el total de resultados posibles.

Ejemplo 1. En una bolsa hay 5 tarjetas numeradas del 1 al 5. Se extrae una al azar. ¿Cuál es la probabilidad de obtener un número par?

S = {1, 2, 3, 4, 5}.
A = {2, 4}; hay 2 resultados favorables.

$$P(A) = 2/5 = 0{,}40.$$

Resultado: 40 %.

Ejemplo 2. Un almacén tiene 20 productos: 8 pequeños, 7 medianos y 5 grandes. Se elige uno al azar. ¿Cuál es la probabilidad de seleccionar uno grande?

Total: n(S) = 20.
Evento A: producto grande; n(A) = 5.

$$P(A) = 5/20 = 0{,}25.$$

Resultado: 25 %.

LIVEWORKSHEETS

Ejemplo 3. Se revisan 50 dispositivos: 38 funcionan correctamente, 7 tienen error de software y 5 presentan daño físico. Se selecciona uno. ¿Cuál es la probabilidad de que tenga algún defecto?

Total: n(S) = 50.
Evento A: dispositivo con defecto.
Favorable: n(A) = 7 + 5 = 12.

$$P(A) = 12/50 = 0{,}24.$$

Resultado: 24 %.

Interpretación: En el lote revisado, aproximadamente 24 de cada 100 dispositivos presentarían defectos si se mantiene el mismo comportamiento.

1.2 Axiomas de probabilidad

Los axiomas son reglas básicas que toda probabilidad debe cumplir. Permiten comprobar que los resultados calculados son coherentes.

- No negatividad: ninguna probabilidad puede ser menor que cero.
- Normalización: la probabilidad del espacio muestral completo es 1.
- Aditividad: si dos eventos no pueden ocurrir simultáneamente, la probabilidad de que ocurra uno u otro es la suma de sus probabilidades.

Se utilizan para verificar cálculos, establecer límites y sumar probabilidades de eventos mutuamente excluyentes.

Fórmulas

$$P(A) \geq 0 \quad ; \quad P(S) = 1 \quad ; \quad P(\emptyset) = 0$$

$$\text{Si } A \cap B = \emptyset, \text{ entonces } P(A \cup B) = P(A) + P(B)$$

Ejemplo 1. Al lanzar un dado, ¿cuál es la probabilidad de obtener un número del 1 al 6?

El evento corresponde a todo el espacio muestral.
P(S) = 1.

Resultado: 1 o 100 %., **Interpretación:** Es un evento seguro.

Ejemplo 2. En una máquina, la probabilidad de que una venta corresponda a agua es 0,30 y a jugo es 0,25. Una venta no puede ser ambas a la vez. ¿Cuál es la probabilidad de que sea agua o jugo?

Los eventos son excluyentes.

$$P(\text{agua o jugo}) = 0{,}30 + 0{,}25 = 0{,}55.$$

Resultado: 0,55 o 55 %.

Ejemplo 3. Un comercio registra que 0,08 de los pedidos se cancelan y 0,12 se devuelven. Para este ejercicio, se consideran categorías excluyentes. ¿Qué probabilidad corresponde a cancelación o devolución?

P(cancelación) = 0,08.
P(devolución) = 0,12.

$$P(C \cup D) = 0{,}08 + 0{,}12 = 0{,}20.$$

Resultado: 20 %.

Interpretación: La probabilidad total de pertenecer a cualquiera de esas dos categorías es 20 %.

LIVEWORKSHEETS

1.3 Álgebra de eventos

El álgebra de eventos permite combinar eventos mediante operaciones de conjuntos:

- Unión ($A \cup B$): ocurre A, ocurre B o ambos.
- Intersección ($A \cap B$): ocurren A y B al mismo tiempo.
- Complemento (A^c): no ocurre A.
- Diferencia ($A - B$): ocurre A, pero no B.

Cuando un problema contiene dos o más condiciones y se necesita identificar si se cumple una, ambas, ninguna o una condición sin la otra.

Fórmulas

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

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

Ejemplo 1. En un grupo, 12 estudiantes usan una aplicación A, 9 usan una aplicación B y 4 usan ambas. ¿Cuántos usan A o B?

Datos:

12 estudiantes usan la aplicación A

9 estudiantes usan la aplicación B

4 estudiantes usan ambos (A Y B)

Aplicar unión: $n(A \cup B) = n(A) + n(B) - n(A \cap B)$.
$n(A \cup B) = 12 + 9 - 4 = 17$.

Resultado: 17 estudiantes.

Ejemplo 2. La probabilidad de que un pedido sea urgente es 0,35. ¿Cuál es la probabilidad de que no sea urgente?

Evento complementario: no urgente.

$P(A^c) = 1 - 0{,}35 = 0{,}65$.

Resultado: 65 %.

Ejemplo 3. En una encuesta de 100 personas, 55 usan plataforma X, 40 usan plataforma Y y 20 usan ambas. ¿Cuál es la probabilidad de que una persona use X o Y?

Datos:

55 personas usan plataforma X

40 personas usan plataforma Y

20 personas usan ambos

$P(X \cup Y) = 55 + 40 - 20 = 75$.
$P(X \cup Y) = 75/100 = 0{,}75$.

Resultado: 75 %.

LIVEWORKSHEETS

EJERCICIOS DE LA UNIDAD 1.

Nombre completo
Fecha:

Desarrollar cada uno de los ejercicios: identifique los datos, aplique la formula y justifique el resultado

1. Espacio muestral: Se lanza un dado de seis caras.

a) Escribe el espacio muestral S.
b) Define A: obtener un número impar.
c) Calcula P(A).

2. Axiomas: En una encuesta, P(A)=0,42. Calcula la probabilidad del complemento y explica por qué no puede ser negativa.

3. Unión: En un grupo de 40 personas, 18 prefieren té, 16 café y 7 ambas bebidas. Calcula cuántas prefieren té o café y la probabilidad correspondiente.

4. Clasificación: Un lote contiene 25 artículos: 10 rojos, 8 azules y 7 verdes. Se selecciona uno al azar. Calcula P(rojo), P(azul) y P(no verde).

5. Aplicación empresarial: De 80 pedidos, 30 son urgentes, 25 tienen envío exprés y 12 cumplen ambas condiciones. Calcula la probabilidad de que un pedido sea urgente o exprés. Interpreta.

LIVEWORKSHEETS

UNIDAD 2. REGLAS DE PROBABILIDAD

2.1 Regla de adición

La regla de adición calcula la probabilidad de que ocurra A o B. Si los eventos pueden coincidir, se resta la intersección para no contarla dos veces. Si son mutuamente excluyentes, la intersección vale cero.

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

Se utiliza cuando el enunciado dice "A o B", "al menos uno" o cuando se busca la probabilidad de pertenecer a una de dos categorías.

Ejemplo 1. La probabilidad de que una persona elija caminar es 0,40 y que elija bicicleta es 0,25. Si las opciones son excluyentes, ¿cuál es la probabilidad de elegir una u otra?

$$P(C \cup B) = 0,40 + 0,25.$$
$$P(C \cup B) = 0,65.$$

Resultado: 65 %.

Ejemplo 2. El 50 % de clientes compra pan, el 35 % compra leche y el 20 % compra ambos. Calcula la probabilidad de que un cliente compre pan o leche.

$$P(P \cup L) = 0,50 + 0,35 - 0,20.$$
$$P(P \cup L) = 0,65.$$

Resultado: 65 %.

Ejemplo 3. En una semana, la probabilidad de una incidencia de acceso es 0,18, de impresión 0,12 y de que ocurran ambas en el mismo registro 0,05. Calcula la probabilidad de que ocurra al menos una.

$$P(A \cup I) = 0,18 + 0,12 - 0,05.$$
$$P(A \cup I) = 0,25.$$

Resultado: 25 %.

Interpretación: Una cuarta parte de los registros podría presentar al menos uno de esos dos tipos de incidencia.

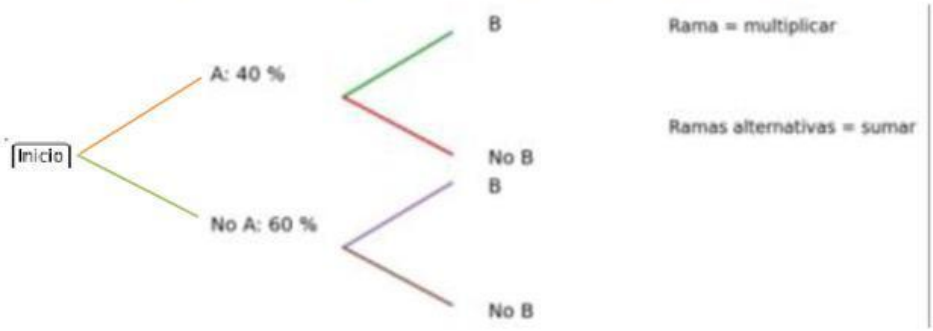
2.2 Regla de multiplicación

La regla de multiplicación calcula la probabilidad de que ocurran dos eventos juntos. Se usa probabilidad condicional para expresar la posibilidad de B cuando ya ocurrió A.

$$P(A \cap B) = P(A) \times P(B|A)$$

Si A y B son independientes, conocer que ocurrió A no cambia la probabilidad de B, por lo que P(B|A)=P(B).

$$\text{Si A y B son independientes: } P(A \cap B) = P(A) \times P(B)$$



LIVEWORKSHEETS

¿Cuándo se utiliza?

- Cuando el problema pide que ocurran dos condiciones a la vez: A y B.
- En procesos por etapas: elegir primero un producto y después otro.
- En selecciones sin reemplazo, cuando la segunda probabilidad cambia después de la primera.

Ejemplo 1. Se lanzan dos monedas juntas. ¿Cuál es la probabilidad de obtener cara en ambas?

Para la primera moneda C1: P(C1)=1/2
Para la segunda moneda C2: P(C2)=1/2
Como son eventos independientes:

$$P(C_1 \cap C_2) = P(C_1) \cdot P(C_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = \mathbf{0.25} \quad (25\%)$$

Resultado: 25 %.

Ejemplo 2. En una caja hay 5 productos buenos y 3 defectuosos. Se extraen dos productos seguidos sin devolver el primero. ¿Cuál es la probabilidad de que ambos sean buenos?

Datos:

- Productos buenos (B): 5
- Productos defectuosos (D): 3

Total de productos: $5 + 3 = 8$

- **Definición de eventos:**
  - B_1: El primer producto extraído es bueno.
  - B_2: El segundo producto extraído es bueno.

Calcular la probabilidad de la primera extracción $\mathbf{P(B_1)}$: En la primera extracción hay 5 productos buenos de un total de 8:

$$P(B_1) = \frac{5}{8}$$

Calcular la probabilidad condicional de la segunda extracción $\boldsymbol{P(B_2 \mid B_1)}$: Dado que el primer producto fue bueno y no se devuelve a la caja, la composición de la caja cambia para la segunda extracción:

- Productos buenos restantes: $5 - 1 = 4$
- Total de productos restantes: $8 - 1 = 7$

Por lo tanto, la probabilidad de sacar un segundo producto bueno es:

$$P(B_2 \mid B_1) = \frac{4}{7}$$

Multiplicamos las probabilidades de ambas etapas:

$$P(B_1 \cap B_2) = P(B_1) \cdot P(B_2 \mid B_1)$$

$$P(B_1 \cap B_2) = \frac{5}{8} \cdot \frac{4}{7} = \frac{20}{56} = 0.35 = 35.71\%$$

Resultado: 35,71 %.

Interpretación: La segunda extracción depende de lo que ocurrió en la primera.

Si el ejercicio dijera *"con devolución"*, los eventos serían independientes y el cálculo sería $\frac{5}{8} \cdot \frac{5}{8} = \frac{25}{64} \approx 39.06\%$.

Al ser *"sin devolución"*, siempre debes restar 1 tanto al número de elementos favorables como al espacio muestral total en la segunda extracción

LIVEWORKSHEETS

Ejemplo 3. Un proceso tiene dos etapas. La probabilidad de que la primera termine correctamente es 0,95. Si termina correctamente, la probabilidad de que la segunda también lo haga es 0,90. Calcula la probabilidad de que ambas terminen correctamente.

Datos:

Evento E_1**:** La primera etapa termina correctamente. $P(E_1) = 0.95$

Evento $E_2 \mid E_1$**:** La segunda etapa termina correctamente, dado que la primera terminó correctamente.

$$P(E_2 \mid E_1) = 0.90$$

Aplicar la Regla de Multiplicación. Sustituimos los datos en la fórmula:

$$P(E_1 \cap E_2) = P(E_1) \cdot P(E_2 \mid E_1)$$

$$P(E_1 \cap E_2) = 0.95 \times 0.90$$

$$P(E_1 \cap E_2) = \mathbf{0.855} \quad (85.5\%)$$

Resultado: La probabilidad de que ambas etapas terminen correctamente es de 85,5 %.

Regla clave: Cuando las etapas ocurren en secuencia y se exige que ocurra la primera **Y** la segunda (intersección de eventos), se **multiplican** las probabilidades.

2.3 Eventos independientes

Dos eventos son independientes cuando la ocurrencia de uno no modifica la probabilidad del otro. No significa que sean eventos distintos o que no puedan ocurrir juntos.

$$\mathbf{P(A \cap B) = P(A) \times P(B)}$$

Se utiliza, por ejemplo, en lanzamientos separados de una moneda o en procesos que no se influyen entre sí. No se debe asumir independencia sin que el contexto lo justifique.

Ejemplo 1. Se lanza un dado y una moneda. ¿Cuál es la probabilidad de obtener 6 y cara?

Datos:

Espacio muestral del dado {1,2,3,4,5,6}, caso favorable de obtener el 6 es 1 caso. P(6)=1/6

Epacio muestral de la moneda {cara, cruz}, caso favorable de obtener cara es 1 caso. P(cara)=1/2.

Como los resultados son independiente, aplicar la regla de multiplicación:

P(6 y cara)=1/6 × 1/2=1/12= 8.33%

Resultado: 8,33 % aproximadamente.

- Usamos **multiplicación** cuando buscamos que ocurra una cosa **Y** la otra al mismo tiempo ($A \cap B$)
- Usamos **suma** cuando buscamos que ocurra una opción **O** la otra ($A \cup B$).

Ejemplo 2. La probabilidad de que un usuario complete correctamente un acceso es 0,8. Para dos intentos independientes, calcula la probabilidad de que ambos sean correctos.

Usuario en el primer acceso P(A)=0,8;

Usuario en el segundo acceso P(B)=0,8.

Los intentos son independientes.

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

$$P(A \cap B) = 0.8 \times 0.8$$

$$P(A \cap B) = \mathbf{0.64} \quad (64\%)$$

Resultado: La probabilidad de que ambos accesos sean correctos es de 0.64 (o 64%).

LIVEWORKSHEETS

Ejemplo 3. La probabilidad de encontrar un artículo defectuoso en el lote A es 0,04 y en el lote B es 0,03. Si los lotes se consideran independientes, calcula la probabilidad de que ambos presenten un artículo defectuoso en las inspecciones.

Datos:

Evento A: Encontrar un artículo defectuoso en el lote A. $P(A) = 0.04$

Evento B: Encontrar un artículo defectuoso en el lote B. $P(B) = 0.03$

Multiplicamos la probabilidad:

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

$$P(A \cap B) = 0.04 \times 0.03$$

$$P(A \cap B) = \mathbf{0.0012} \quad (0.12\%)$$

Resultado: 0,12 %. La probabilidad de que ambos lotes presenten un artículo defectuoso en las inspecciones es de 0.0012 (o 0.12%).

Recuerda hacer el conteo de decimales: 0.04 tiene 2 decimales y 0.03 tiene 2 decimales, por lo que el resultado debe tener $2 + 2 = 4$ lugares decimales (0.0012).

2.4 Probabilidad condicional

La probabilidad condicional mide la posibilidad de que ocurra B sabiendo que A ya ocurrió. El conjunto de información disponible se reduce a los casos en que A ocurrió.

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

con P(B) > 0

Enfoque de Espacio Muestral Reducido (Regla de Laplace): Cuando trabajamos con conteo directo de personas o elementos:

$$P(A \mid B) = \frac{\textbf{Número de elementos que cumplen } A \textbf{ dentro del grupo } B}{\textbf{Número total de elementos del nuevo universo } B} = \frac{|A \cap B|}{|B|}$$

Se utiliza cuando el enunciado incluye expresiones como “sabiendo que”, “dado que”, “entre los que cumplen” o “si ya ocurrió”.

Ejemplo 1. En un curso, 20 estudiantes aprobaron un examen y 12 de ellos entregaron el proyecto. Si se elige a alguien entre quienes aprobaron, ¿cuál es la probabilidad de que haya entregado el proyecto?

Datos:

Evento A**:** El estudiante aprobó el examen.
Evento B**:** El estudiante entregó el proyecto.

Estudiantes que aprobaron el examen: $|A| = 20$
Estudiantes que aprobaron el examen **Y** entregaron el proyecto: $|A \cap B| = 12$

$$P(\text{Entregó proyecto} \mid \text{Aprobó examen}) = \frac{|A \cap B|}{|A|}$$

$$P(B \mid A) = \frac{12}{20} = 0.60$$

Resultado: 60 %.

LIVEWORKSHEETS

Ejemplo 2. De 40 productos, 10 son defectuosos y 6 de esos defectuosos tienen daño en el empaque. Si se sabe que el producto es defectuoso, calcula la probabilidad de daño en el empaque.

Total de productos: 40

Evento D: El producto es defectuoso. Cantidad de productos defectuosos: $|D| = 10$

Evento E: El producto tiene daño en el empaque.

Evento $D \cap E$: El producto es defectuoso Y tiene daño en el empaque.

Cantidad de productos defectuosos con daño en el empaque: $|D \cap E| = 6$

$$P(E \mid D) = \frac{|D \cap E|}{|D|}$$

$$P(E \mid D) = \frac{6}{10} = 0.60$$

Resultado: 60 %.

Ejemplo 3. En 200 accesos, 50 fueron bloqueados y 30 de esos bloqueos se debieron a contraseña incorrecta. Si un acceso fue bloqueado, ¿cuál es la probabilidad de que la causa sea contraseña incorrecta?

Total de accesos registrados: 200

Evento B: El acceso fue bloqueado. Cantidad de accesos bloqueados: $|B| = 50$

Evento C: El bloqueo se debió a contraseña incorrecta.

Evento $B \cap C$: El acceso fue bloqueado Y la causa fue contraseña incorrecta.

Cantidad de accesos bloqueados por contraseña **incorrecta:** $|\boldsymbol{B \cap C}| = \mathbf{30}$

$$P(C \mid B) = \frac{|B \cap C|}{|B|}$$

$$P(C \mid B) = \frac{30}{50} = 0.60$$

Resultado: 60 %.

Interpretación: La respuesta no se calcula sobre 200, porque la pregunta limita el análisis a los accesos bloqueados.

LIVEWORKSHEETS

Ejercicios de la Unidad 2

Nombre completo:
Fecha:

Desarrollar cada uno de los ejercicios: identifique los datos, aplique la formula y expresa el resultado en porcentaje.

1. P(A)=0,45; P(B)=0,30 y P(A∩B)=0,10. Calcula P(AUB).

2 Una bolsa tiene 4 fichas rojas y 6 azules. Se extraen dos sin reemplazo. Calcula la probabilidad de que ambas sean rojas.

3. La probabilidad de que un correo sea spam es 0,20 y la de que un archivo adjunto esté dañado es 0,05. Suponiendo independencia, calcula la probabilidad de que ocurran ambos eventos.

4. De 60 pedidos entregados, 18 llegaron con retraso. Si se elige un pedido entre los entregados, ¿cuál es la probabilidad de que haya llegado con retraso?

5. En una empresa, 35 % de usuarios usa el sistema móvil, 50 % usa el sistema web y 20 % usa ambos. Calcula la probabilidad de usar al menos uno. Luego, si 100 usuarios usan el móvil y 40 de ellos también usan web, calcula P(web|móvil).

LIVEWORKSHEETS

UNIDAD 3. TEOREMAS DE PROBABILIDAD

3.1 Teorema de Probabilidad Total

Se utiliza cuando un resultado puede provenir de varios grupos o causas que dividen el total. La probabilidad del resultado se obtiene sumando las probabilidades de cada ruta posible.

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots + P(B|A_n)P(A_n)$$

Los grupos A_1, A_2, … deben ser mutuamente excluyentes y cubrir todo el espacio muestral.

- Cuando los productos proceden de diferentes proveedores o líneas.
- Cuando los usuarios pertenecen a distintos canales o segmentos.
- Cuando una falla puede originarse en varias áreas o procesos.

Ejemplo 1. Una tienda compra 60 % de sus productos al proveedor A y 40 % al proveedor B. La tasa de defecto es 2 % para A y 5 % para B. Calcula la probabilidad de que un producto elegido sea defectuoso.

Datos:

Evento A: El producto proviene del proveedor A. $P(A) = 60\% = 0.60$

Evento B: El producto proviene del proveedor B. $P(B) = 40\% = 0.40$

Evento D: El producto elegido es defectuoso.

Tasas de defecto (probabilidades condicionales):

- $P(D \mid A) = 2\% = 0.02$ (Probabilidad de que sea defectuoso dado que es de A)
- $P(D \mid B) = 5\% = 0.05$ (Probabilidad de que sea defectuoso dado que es de B)

$$P(D) = P(A) \cdot P(D \mid A) + P(B) \cdot P(D \mid B)$$
$$P(D) = (0.60 \times 0.02) + (0.40 \times 0.05)$$
$$P(D) = 0.012 + 0.020 = \mathbf{0.032} \quad (3.2\%)$$

Resultado: La probabilidad de que un producto elegido al azar sea defectuoso es de 0.032 (o 3.2%).

Ejemplo 2. El 70 % de consultas llega por web y 30 % por aplicación. El 4 % de las consultas web y el 8 % de las consultas por aplicación tienen datos incompletos. Calcula la probabilidad total de recibir una consulta incompleta.

- **Evento** W: La consulta llega por la Web. $P(W) = 70\% = 0.70$
- **Evento** A: La consulta llega por la Aplicación. $P(A) = 30\% = 0.30$
- **Evento** I: La consulta tiene datos incompletos.

$P(I \mid W) = 4\% = 0.04$ (Probabilidad de estar incompleta dado que proviene de la Web)

$P(I \mid A) = 8\% = 0.08$ (Probabilidad de estar incompleta dado que proviene de la Aplicación)

fórmula de Probabilidad Total:

$$P(I) = P(W) \cdot P(I \mid W) + P(A) \cdot P(I \mid A)$$
$$P(I) = (0.70 \times 0.04) + (0.30 \times 0.08) = 0.052$$

Resultado: La probabilidad total de recibir una consulta incompleta es de 0.052 (o 5.2%).

LIVEWORKSHEETS

Ejemplo 3. Una empresa recibe incidencias de tres áreas: ventas (50 %), soporte (30 %) y logística (20 %). Las tasas de registro incorrecto son 1 %, 3 % y 5 %, respectivamente. Calcula la probabilidad de que una incidencia elegida tenga registro incorrecto.

Definir los eventos:

- **Evento** V: La incidencia proviene del área de Ventas $P(V) = 50\% = 0.50$
- **Evento** S: La incidencia proviene del área de Soporte. $P(S) = 30\% = 0.30$
- **Evento** L: La incidencia proviene del área de Logística. $P(L) = 20\% = 0.20$
- **Evento** I: La incidencia tiene un registro incorrecto.

Tasas de registro:

- $P(I \mid V) = 1\% = 0.01$ (Probabilidad de registro incorrecto en Ventas)
- $P(I \mid S) = 3\% = 0.03$ (Probabilidad de registro incorrecto en Soporte)
- $P(I \mid L) = 5\% = 0.05$ (Probabilidad de registro incorrecto en Logística)

$$P(I) = P(V) \cdot P(I \mid V) + P(S) \cdot P(I \mid S) + P(L) \cdot P(I \mid L)$$

$$P(I) = (0.50 \times 0.01) + (0.30 \times 0.03) + (0.20 \times 0.05) = 0.024$$

Resultado: La probabilidad de que una incidencia elegida al azar tenga un registro incorrecto es de 0.024 (o 2.4%).

Interpretación: Se pondera la tasa de error de cada área por la proporción de incidencias que aporta.

3.2 Teorema de Bayes

El teorema de Bayes permite actualizar una probabilidad cuando se conoce nueva información. Responde preguntas inversas: si se observó un resultado, ¿qué probabilidad hay de que provenga de una causa determinada?

$$\boldsymbol{P(A_i \mid B) = \frac{P(A_i) \cdot P(B \mid A_i)}{P(B)}}$$

P(A) es la probabilidad inicial; P(B|A) es la probabilidad de observar B cuando A ocurre; P(B) es la probabilidad total de observar B.

Teorema de Bayes con la probabilidad total: Dado que el evento B puede ocurrir a través de distintas causas o particiones $A_1, A_2, \ldots, A_n$, expandimos el denominador usando el Teorema de la Probabilidad Total:

$$\boldsymbol{P(A_i \mid B) = \frac{P(A_i) \cdot P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j) \cdot P(B \mid A_j)}}$$

- $P(A_i)$: Probabilidad a Priori (probabilidad inicial de la causa A_i antes de conocer la nueva evidencia).
- $P(B \mid A_i)$: Verosimilitud (probabilidad de que ocurra el evento B si la causa A_i es verdadera).
- $P(B)$: Probabilidad Total de la evidencia (la probabilidad global de que el evento B suceda).
- $P(A_i \mid B)$: Probabilidad a Posteriori (la nueva probabilidad ajustada de la causa A_i tras confirmar que ocurrió B).

¿Cuándo se utiliza?

- Para identificar el proveedor más probable de un producto defectuoso.
- Para estimar la causa probable de una alerta o error.
- Para interpretar resultados de filtros, clasificadores y pruebas de detección.

LIVEWORKSHEETS

Ejemplo 1. El 60 % de productos viene de A y el 40 % de B. A tiene 2 % de defectos y B 5 %. Si se encontró un producto defectuoso, calcula la probabilidad de que provenga de B.

Paso 1: identificación de eventos:

Probabilidades de origen (A priori): $P(A) = 60\% = 0.60$, $P(B) = 40\% = 0.40$

Tasas de defecto (Verosimilitud / Probabilidades condicionales):

- $P(D \mid A) = 2\% = 0.02$ (Defectuoso dado que es del proveedor A)
- $P(D \mid B) = 5\% = 0.05$ (Defectuoso dado que es del proveedor B)

Paso 2: Calcular la Probabilidad Total del evento observado $P(D)$.

$$P(D) = P(A) \cdot P(D \mid A) + P(B) \cdot P(D \mid B)$$
$$P(D) = (0.60 \times 0.02) + (0.40 \times 0.05)$$
$$P(D) = 0.012 + 0.020 = \mathbf{0.032} \quad (3.2\%)$$

Paso 3: Aplicar el Teorema de Bayes.

$$P(B \mid D) = \frac{P(B) \cdot P(D \mid B)}{P(D)}$$

$$P(B \mid D) = \frac{0.40 \times 0.05}{0.032}$$

$$P(B \mid D) = \frac{0.020}{0.032} = \frac{20}{32} = 0.62$$

Resultado: 62 %. Sabiendo que el producto inspeccionado resultó ser defectuoso, la probabilidad de que provenga del proveedor B es de

Ejemplo 2. El 30 % de correos es spam. Un filtro marca correctamente el 90 % del spam y marca por error el 10 % de correos legítimos. Si un correo fue marcado, ¿cuál es la probabilidad de que realmente sea spam?

Paso 1: Identificar los eventos y las probabilidades iniciales.

- **Evento** S: El correo es realmente Spam. $P(S) = 30\% = 0.30$
- **Evento** L: El correo es Legítimo (No Spam). $P(L) = 100\% - 30\% = 70\% = 0.70$
- **Evento** M: El filtro marca el correo.

Paso 2. Probabilidades condicionales (Comportamiento del filtro):

- $P(M \mid S) = 90\% = 0.90$ (Verdadero Positivo: detecta el Spam)
- $P(M \mid L) = 10\% = 0.10$ (Falso Positivo: marca por error un correo legítimo)

$$P(M) = P(S) \cdot P(M \mid S) + P(L) \cdot P(M \mid L)$$

$$P(M) = (0.30 \times 0.90) + (0.70 \times 0.10) = 0.34$$

Paso 3. Aplicar el Teorema de Bayes. Calculamos la probabilidad a posteriori de que sea **Spam** dado que fue **Marcado** ($P(S \mid M)$):

$$P(S \mid M) = \frac{P(S) \cdot P(M \mid S)}{P(M)}$$

$$P(S \mid M) = \frac{0.30 \times 0.90}{0.34} = \frac{0.27}{0.34} = 0.79$$

Resultado: Si un correo fue marcado por el filtro, la probabilidad de que realmente sea spam es de aproximadamente 0.79 (o 79%).

LIVEWORKSHEETS

Ejemplo 3. Una fábrica produce 50 % de unidades en la línea A y 50 % en la línea B. El clasificador detecta como defectuosas el 95 % de las unidades realmente defectuosas de A y el 90 % de las de B. La tasa de defectos es 4 % en A y 6 % en B. Si el sistema detecta una unidad como defectuosa, calcula la probabilidad de que proceda de B.

3. RESOLUCIÓN PASO A PASO DEL EJERCICIO

Paso 1: Identificar las probabilidades iniciales y definir los eventos.

Eventos de Producción:	Eventos de Falla Real	(Sensibilidad):
$P(A) = 50\% = 0.50$ (Producción en Línea A)	$P(\text{Defecto Real en A}) = 4\% = 0.04$	Detecta el 95% de los defectos reales de A (0.95).
$P(B) = 50\% = 0.50$ (Producción en Línea B)	$P(\text{Defecto Real en B}) = 6\% = 0.06$	Detecta el 90% de los defectos reales de B (0.90).
Probabilidad conjunta de que una unidad sea detectada como defectuosa (D) para cada línea:		
Para la Línea A: $P(D \mid A) = 0.04 \times 0.95 = 0.038$ (3.8% de toda la producción de A es detectada como defectuosa) Para la Línea B: $P(D \mid B) = 0.06 \times 0.90 = 0.054$ (5.4% de toda la producción de B es detectada como defectuosa)		

Paso 2: Calcular la Probabilidad Total del evento observado $P(D)$

$$P(D) = P(A) \cdot P(D \mid A) + P(B) \cdot P(D \mid B)$$

$$P(D) = (0.50 \times 0.038) + (0.50 \times 0.054) = 0.046$$

Paso 3: Aplicar el Teorema de Bayes. Calculamos la probabilidad a posteriori de que la unidad proceda de la **Línea B** dado que el sistema la detectó como **defectuosa** $(P(B \mid D))$:

$$P(B \mid D) = \frac{P(B) \cdot P(D \mid B)}{P(D)}$$

$$P(B \mid D) = \frac{0.50 \times 0.054}{0.046} = \frac{0.027}{0.046} = 0.586$$

Resultado: 0.5870 (o 58.70%)

Interpretación: Entre las unidades marcadas como defectuosas, cerca del 58,7 % procede de la línea B.

3.3 Filtros de clasificación

Un filtro de clasificación asigna elementos a categorías a partir de características o reglas.

Un **filtro de clasificación** (o *clasificador*) es un sistema, algoritmo o prueba diagnóstica diseñado para tomar un conjunto de datos o elementos y asignarlos a una categoría específica, generalmente evaluando si presentan o no cierta característica de interés (como *Defectuoso / Bueno*, *Spam / No Spam* o *Enfermo / Sano*).

	Condición Real: POSITIVA (Tiene la falla / Es Spam)	**Condición Real: NEGATIVA** (Está sano / Es Limpio)
Filtro Detecta: POSITIVO	Verdadero Positivo (VP) *(Acierto: Atrapó la falla)*	Falso Positivo (FP) *(Error: Falsa alarma)*
Filtro Detecta: NEGATIVO	Falso Negativo (FN) *(Error: Se le escapó la falla)*	Verdadero Negativo (VN) *(Acierto: Dejó pasar lo bueno)*

LIVEWORKSHEETS

Ejemplos: correo spam/no spam, transacciones sospechosas/no sospechosas, productos conformes/no conformes o registros completos/incompletos.

- **Verdadero positivo**: el elemento es positivo y el filtro lo identifica como positivo.
- **Falso positivo:** el elemento es negativo, pero el filtro lo marca como positivo.
- **Verdadero negativo:** el elemento es negativo y el filtro lo identifica como negativo.
- **Falso negativo:** el elemento es positivo, pero el filtro no lo detecta.

La probabilidad total permite calcular cuántos elementos serán marcados; Bayes permite calcular qué proporción de los marcados realmente pertenece a la categoría buscada.

Ejemplo 1. De 100 mensajes, 20 son spam. Un filtro identifica 18 de esos 20 como spam. ¿Cuántos verdaderos positivos detecta?

Verdaderos positivos = mensajes spam correctamente identificados.

categoría	Condición Real	Resultado del Filtro	Cantidad
Verdaderos Positivos (VP)	Spam	Marcado como Spam	18
Falsos Negativos (FN)	Spam	Marcado como No Spam (se le escapó)	2
Verdaderos Negativos (VN)	No Spam (80)	Marcado como No Spam	80 *
Falsos Positivos (FP)	No Spam (80)	Marcado como Spam	0 *

Resultado: 18 verdaderos positivos.

Ejemplo 2. De 200 transacciones, 10 son sospechosas. Un sistema alerta sobre 8 sospechosas y además alerta por error sobre 18 normales. ¿Cuántas alertas son correctas y qué proporción representan entre todas las alertas?

Desglose de las transacciones (Matriz de confusión)

- Total de transacciones: 200
  - Transacciones Sospechosas: 10
  - Transacciones Normales: $200 - 10 = 190$
- Comportamiento del sistema de alertas
  - Alertas sobre sospechosas (Verdaderos Positivos - VP): 8
  - Alertas por error sobre normales (Falsos Positivos - FP): 18
  - Total de alertas emitidas por el sistema: Total de alertas $= 8 + 18 = 26$
  -

¿Cuántas alertas son correctas? Son **8 alertas correctas** (los Verdaderos Positivos).

¿Qué proporción representan entre todas las alertas? Esta métrica se conoce como **Precisión** (o *Valor Predictivo Positivo*):

$$\text{Proporción} = \frac{\text{Alertas correctas}}{\text{Total de alertas}} = \frac{8}{26} = \frac{4}{13} \approx 0.3077$$

Las alertas correctas son **8** y representan el 30.77% del total de alertas emitidas por el sistema.

LIVEWORKSHEETS