
MAGNA EDUCATIO AD MELIOR
UNIVERSIDAD DE SANTANDER
Universidad de Santander
UDES
VIGILADA MINEDUCACIÓN | SNIES 2832

Estadística

Probabilidad

Daniel Martínez Bello

Universidad de Santander
Maestría en Biotecnología

Octubre de 2025

Experimentos

Conjuntos

Probabilidad

Experimentos

Experimento es un operación planificada bajo condiciones controladas. Si los resultados de un experimento no están predeterminados, se dice que el experimento es un experimento aleatorio, o un experimento probabilístico. Los resultados de un experimento son:

- Un conjunto de mediciones de una población de la cual se ha obtenido una muestra al azar.
- Mediciones de un ensayo realizado en un laboratorio.
- Resultados de un ensayo clínico aleatorizado
- Resultados de una simulación de un proceso tecnológico
- Los indicadores clínicos de las historias clínicas de pacientes intervenidos con un tratamiento en una institución de salud

Algunas definiciones intuitivas sobre la probabilidad

- Probabilidad subjetiva: la probabilidad como medida de incertidumbre definida por el observador
- Probabilidad empírica: probabilidad derivada de un muestreo sobre una población
- Probabilidad basada en frecuencias.

$$P(E_k) = \frac{E_k}{\sum_k E}$$

Notación

- El **espacio muestral**, Ω, es el conjunto de posibles resultados de un experimento
- Un **evento**, denotado como E, es un conjunto de Ω

  Ejemplo: lanzar un dado, y registrar los números pares $E = \{2, 4, 6\}$
- Un evento **elemental** es un resultado particular del experimento

  Ejemplo: lanzar un dado y que sea un cuatro, $\omega = 4$
- $\emptyset$ se conoce como el **evento núlo** o el **conjunto vacio**

Ideas básicas

Estrategía utilizada en la ciencia:

- Para un experimento probabilístico dado:
  - Asigne todo lo que se conoce a o se teoriza a un modelo sistemático (función matemática)
  - Asigne todo lo demas al azar, *aun si el proceso bajo estudio no necesariamente es aleatorio*
  - Utilice la probabilidad para cuantificar la incertidumbre en sus conclusiones
  - Evalue la sensibilidad de sus conclusiones a los supuestos de su modelo

Medida de probabilidad

Una **medida de probabilidad**, P, es una función real del conjunto de eventos posibles que obedecen los siguientes axiomas

1. Para un evento $E \subset \Omega$, $0 \leq P(E) \leq 1$
2. $P(\Omega) = 1$
3. Si dos eventos E_1 y E_2 son mutualmente excluyentes, esto es $(E_1 \cap E_2 = \emptyset)$, entonces $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

Regla de la adición para dos eventos (caso general)

$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

Nota

- P se define en $\mathcal{F}$ como una colección de subconjuntos de Ω
- Ejemplo $\Omega = \{1, 2, 3\}$ luego

$$\mathcal{F} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

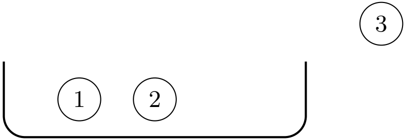
- Cuando Ω es un conjunto continuo, la definición es mas complicada. En este caso, se asume que $\mathcal{F}$ es suficientemente denso de forma que cualquier conjunto en el que estemos interesados estará en el.

De las definiciones anteriores se desprende que:

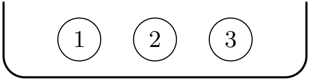
- $P(\emptyset) = 0$
- $P(E) = 1 - P(E^c)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Muestreo con reemplazo y sin reemplazo

La diferencia entre muestreo con reemplazo y sin reemplazo. En el primer muestreo, supongamos que se hace al azar y se extrae la ficha 3

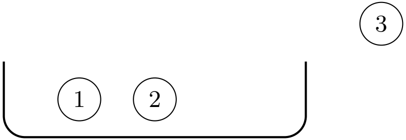


Muestreo con reemplazo, el segundo muestreo se lleva a cabo sobre

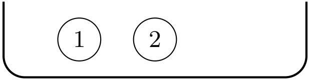


Muestreo con reemplazo y sin reemplazo

La diferencia entre muestreo con reemplazo y sin reemplazo. En el primer muestreo, supongamos que se hace al azar y se extrae la ficha 3



Muestreo sin reemplazo, el segundo muestreo se lleva a cabo sobre

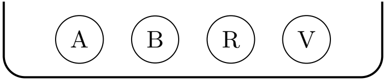


Cuando se hace un muestreo al azar, todas las fichas en la caja tienen la misma oportunidad de ser muestreadas.

Probabilidad condicional

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

Una urna tiene cuatro fichas, amarilla (A), blanca (B), roja (R), verde (V)



Se seleccionan al azar dos fichas, sin reemplazamiento. Cual es la posibilidad de obtener una ficha roja, y luego una blanca?
En el primer muestreo, la posibilidad de roja es uno de cuatro $(1/4)$, y luego en el segundo muestreo la posibilidad de blanca es uno de tres $(1/3)$,luego la posibilidad es

$$\frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$$

- Ejemplo: se realiza un muestreo al azar sobre una población, para establecer la asociación entre el género de esta población y el hábito de fumar

Género	Fumar	Género	Fumar
1	0	0	1
1	1	1	0
0	1	0	0
1	1	0	1
0	1	1	0
0	0	1	0
0	1	0	1
1	1	0	0
1	1	1	1
1	1	1	1
1	0	1	1
0	1	1	1
1	1	1	1
0	0	0	1
0	0	0	1
0	1	0	1
0	1	0	0
0	1	0	1
0	1	1	1
1	1	0	1
0	1	1	1
1	1	1	1
0	1	1	1
0	0	0	1

	Fumar	
Sexo	No	Si
Masculino	8	20
Femenino	5	17

	Fumar		
Sexo	No	Si	Marginal filas
Masculino	8	20	28
Femenino	5	17	22
Marginal Columnas	13	37	50

	Fumar		
Sexo	No	Si	Marginal filas
Masculino	0.16	0.40	0.56
Femenino	0.10	0.34	0.44
Marginal Columnas	0.26	0.74	1

- Cual es la probabilidad de seleccionar una persona al azar de esta población y que sea hombre y no fume.
- Cual es la probabilidad de seleccionar una persona al azar de esta población y que sea mujer y no fume.
- Cuál es la probabilidad de seleccionar una persona al azar de esta población que sea hombre, dado que no fuma.
- Cuál es la probabilidad de seleccionar una persona al azar, que sea mujer dado que no fuma.
- Cuál es la probabilidad de seleccionar una persona al azar que sea mujer dado que no fuma.
- Cuál es la probabilidad de seleccionar una persona al azar, que sea mujer dado que fuma.

Regla de la multiplicación, para el caso general

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

- La posibilidad que dos eventos ocurran es igual a la posibilidad que el primero ocurra, multiplicado por la posibilidad que el segundo ocurra dado que el primero ha ocurrido.
- Ejemplo: en una baraja inglesa, se dispone de 52 cartas, organizadas en corazones $\heartsuit$, diamantes $\diamondsuit$, picas $\spadesuit$ y treboles $\clubsuit$. Cada grupo tiene diez cartas numeradas del uno al diez, y tres cartas que son el principe (J), la reina (Q) y el rey (K.) Se baraja un mazo de cartas, y se reparten dos cartas: Cúal es la posibilidad que la primera carta sea el $\boxed{7\ \heartsuit}$ y la segunda sea la $\boxed{Q\ \spadesuit}$
-

$$P(\boxed{7\ \heartsuit}) = 1/52$$

$$P(\boxed{Q\ \spadesuit} | \boxed{7\ \heartsuit}) = 1/51$$

$$P(\boxed{Q\ \spadesuit} \cap \boxed{7\ \heartsuit}) = P(\boxed{Q\ \spadesuit})P(\boxed{Q\ \spadesuit} | \boxed{7\ \heartsuit}) = 1/52 \times 1/51 = 1/2652$$

Regla de Bayes

Ley de la probabilidad total

Si $E_1, E_2, \cdots, E_n$ son eventos disjuntos uno a uno, que cubren el espacio muestral (Ω) completo (forman una partición de Ω), se puede calcular la probabilidad de cualquier evento A utilizando:

$$P(A) = \sum_i P(A|E_i)P(E_i)$$

Ley de la probabilidad total

Una encuesta clasifica personas según el estrato socioeconómico y el grupo político. El 80% clasifica como estrato 1 a 3, y el 20% como estrato 4 a 6. El 70% de los de estrato 1 a 3 vota independiente, mientras que el 40% de los de estrato 4 a 6 vota independiente.
¿Cuál es la probabilidad de que una persona elegida al azar vote independiente?

- P(Est 1-3) = 0.8
- P(Est 4-6) = 0.2
- P(Ind|Est 1-3) = 0.7
- P(Ind|Est 4-6) = 0.4

$$
\begin{aligned}
P(\text{Ind}) &= P(\text{Ind}|\text{Est 1-3})P(\text{Est 1-3}) + P(\text{Ind}|\text{Est 4-6})P(\text{Est 4-6}) \\
&= 0.7 \times 0.8 + 0.4 \times 0.2 = 0.64
\end{aligned}
$$

Regla de Bayes

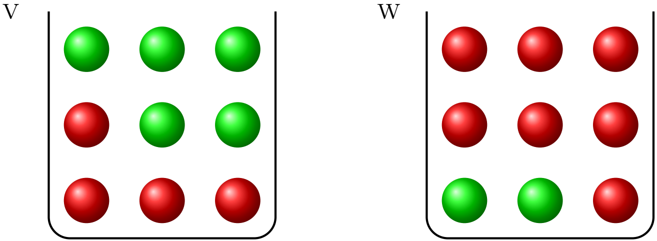
Formulación

$$P(E_k|A) = \frac{P(A|E_k)P(E_k)}{\sum_i P(A|E_i)P(E_i)}$$

para cualquier k definido $(k = 1, 2, \cdots, n)$ Las $P(E_i)$ se conocen como probabilidades previas, y las $P(E_k|A)$ son probabilidades posteriores.

Regla de Bayes

Dos urnas idénticas contienen 9 pelotas. La urna V contiene 4 bolas rojas y 5 bolas verdes. La urna W contiene 7 bolas rojas y dos bolas verdes.



Seleccione una urna al azar, y sin mirar su contenido adivine si es la urna V o W. La probabilidad de seleccionar la urna V es igual a la probabilidad de la urna W y es igual a 1/2.

$$P(W) = P(V) = 1/2$$

Regla de Bayes

Ahora, le permiten extraer una pelota al azar de la urna seleccionada, y mirar su color. Suponga que la pelota extraída es de color verde. Calcule la probabilidad que la bola viene de la la urna V?

$$
\begin{aligned}
P(V|\text{verde}) &= \frac{P(\text{verde}|V)P(V)}{P(\text{verde}|V)P(V) + P(\text{verde}|W)P(W)} \\
&= \frac{5/9 \times 1/2}{5/9 \times 1/2 + 2/9 \times 1/2} = 5/7 = 0.71
\end{aligned}
$$

$$
\begin{aligned}
P(W|\text{verde}) &= \frac{P(\text{verde}|W)P(W)}{P(\text{verde}|V)P(V) + P(\text{verde}|W)P(W)} \\
&= \frac{2/9 \times 1/2}{5/9 \times 1/2 + 2/9 \times 1/2} = 2/7 = 0.29
\end{aligned}
$$

Pruebas diagnósticas

		Enfermedad en el individuo Positivo	Enfermedad en el individuo Negativo	
Resultado prueba diagnóstica	Positivo	n_{11}	n_{12}	n_{1+}
	Negativo	n_{21}	n_{22}	n_{2+}
		n_{+1}	n_{+2}	n_{++}

- T+ = la prueba diagnóstica es positiva.
- T- = la prueba diagnóstica es negativa.
- Z+ = el individuo tiene la enfermedad.
- Z- = el individuo no tiene la enfermedad.

Regla de Bayes Probabilidades condicionales, que establecen la sensibilidad, especificidad, la probabilidad de falsos negativos y de falso positivo de la prueba diagnóstica:

- $P(T+|Z+) = \frac{n_{11}}{n_{+1}}$ = la sensibilidad de la prueba diagnóstica.
- $P(T-|Z+) = \frac{n_{21}}{n_{+2}}$ = la probabilidad de un falso negativo.
- $P(T-|Z-) = \frac{n_{22}}{n_{+2}}$ = la especificidad de la prueba diagnóstica.
- $P(T+|Z-) = \frac{n_{12}}{n_{+1}}$ = la probabilidad de un falso positivo.
- $P(Z+) = \frac{\text{Individuos positivos}}{\text{Total de individuos}}$ = la prevalencia de la enfermedad (determinada por muestreos previos aleatorios en la población).

Regla de Bayes Una prueba diagnóstica para detectar VIH por ELISA, se prueba en 1980 personas, con los siguientes resultados:

		Enfermedad en el individuo	
		Positivo	Negativo
Resultado prueba diagnóstica	Positivo	890	45
	Negativo	15	1030
		905	1075

Regla de Bayes

$$
\begin{aligned}
P(T+|Z+) &= \frac{890}{905} = 0.98 \\
P(T-|Z+) &= \frac{15}{905} = 0.02 \\
P(T-|Z-) &= \frac{1030}{1075} = 0.96 \\
P(T+|Z-) &= \frac{45}{1075} = 0.04 \\
P(Z+) &= \frac{250}{18000} = 0.014
\end{aligned}
$$

La sensibilidad es igual a 0.98, la especificidad es igual a 0.96, y la prevalencia es igual a 0.014

Regla de Bayes ¿Cuál es la probabilidad de que una persona diagnosticada positiva, efectivamente sea positiva?

$$P(Z+|T+) = \frac{P(T+|Z+)P(Z+)}{P(T+|Z+)P(Z+) + P(T+|Z-)P(Z-)}$$
$$= \frac{0.98 \times 0.014}{0.98 \times 0.014 + 0.04 \times 0.986} = 0.258$$

¿Cuál es la probabilidad de que una persona diagnosticada negativa, sea positiva?

$$P(Z+|T-) = \frac{P(T-|Z+)P(Z+)}{P(T-|Z+)P(Z+) + P(T-|Z-)P(Z-)}$$
$$= \frac{0.02 \times 0.014}{0.02 \times 0.014 + 0.96 \times 0.986} = 0.0003$$

Regla de Bayes

Conclusión

- La regla de Bayes permite actualizar el conocimiento actual con el conocimiento previo, de tal forma que se obtiene un conocimiento posterior.
- Es la base para incorporar información externa dentro de un sistema de toma de decisión.

Regla de Bayes

Ejercicio en clase

Del total de accidentes ocurridos en un año en vehículos livianos, el 40% ocurren en camionetas, y el 60% ocurren en automoviles. De cada 100.000 accidentes, 20 involucran fatalidades en automoviles, y 25 involucran fatalidades en camionetas. ¿Determine la probabilidad de que un accidente con fatalidad involucre una camioneta?

Regla de Bayes

Solución

- P(Auto) = 0.6
- P(Camioneta) = 0.4
- P(fatal|Auto) = 20/100.000 = 0.0002
- P(fatal|Camioneta) = 25/100.000 = 0.00025

$$P(\text{Cam}|\text{fatal}) = \frac{P(\text{fatal}|\text{Cam})P(\text{Cam})}{P(\text{fatal}|\text{Cam})P(\text{Cam}) + P(|\text{Auto})P(\text{Auto})}$$
$$= \frac{0.4 \times 0.00025}{0.4 \times 0.00025 + 0.6 \times 0.00020} = 0.54$$

Regla de Bayes

Ejercicio

En una gasolinera, 40% de los clientes utilizan gasolina regular (A1), 35% usan gasolina plus (A2) y 25% utilizan premium (A3). De los clientes que utilizan gasolina regular, sólo 30% llenan sus tanques (evento B). De los clientes que utilizan plus, 60% llenan sus tanques, mientras que los que utilizan premium, 50% llenan sus tanques. Si un cliente llena el tanque, ¿cuál es la probabilidad que pida gasolina regular? ¿Plus? ¿Premium?[1]

[1] Fuente: Devore. Probabilidad y estadística para ingenieros, edición 8, 2012

Regla de Bayes

Ejercicio

Una compañía que vende celulares promociona un modelo básico y un modelo de lujo. Durante el año pasado, 40% de los celulares vendidos fueron del modelo básico. De aquellos que compraron el modelo básico, 30% adquirieron una garantía ampliada, en tanto que 50% de los que compraron el modelo de lujo también lo hicieron. Si se sabe que un comprador seleccionado al azar tiene una garantía ampliada, ¿qué tan probable es que él o ella tengan un modelo básico?[2]

[2] Fuente: Devore. Probabilidad y estadística para ingenieros, edición 8, 2012

Regla de Bayes

Ejercicio

Un gran operador de complejos de tiempo compartido requiere que cualquier persona interesada en hacer una compra primero visite el sitio de interés. Los datos históricos indican que el 20% de todos los compradores potenciales seleccionaron un día de visita, el 50% elige una visita de una noche y el 30% opta por una visita de dos noches. Además, el 10% de los visitantes de un día en última instancia, hacen una compra, el 30% de los visitantes de una noche compran una unidad y el 20% de los visitantes de dos noches deciden comprar. Supongamos que un visitante es seleccionado al azar y se demuestra que ha realizado una compra. ¿Qué tan probable es que esta persona haya realizado una visita de día? ¿Una visita de una noche? ¿Una visita de dos noches?[3]

[3] Fuente: Devore. Probabilidad y estadística para ingenieros, edición 8, 2012