Probabilidad y Estadística I
Semana 3

Teorema de probabilidad total, teorema de Bayes e independencia + Introducción a variables aleatorias

Profesor: Nicolás López MSc

Universidad del Rosario

Contenido

Universidad del Rosario

Probabilidad condicional

Sean A y B dos eventos. Se define la probabilidad del evento A bajo la condición B como

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

Esta última igualdad al multiplicar por $P(B)$ en ambos lados puede verse como

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

Eventos independientes

Definición

Algunas veces, el conocimiento de B no afecta la probabilidad de ocurrencia de un evento A, es decir, $P(A|B) = P(A)$, con lo cual

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

Eventos independientes

Sean A y B dos eventos. Se dice que los eventos son independientes si y solo si

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

Si la igualdad no se cumple, los eventos son **dependientes**.

Ejemplo. Hipertensos y fumadores

Se sabe que en una población el chance de ser hipertenso es 0.2, mientras que el de ser fumador es 0.45. Además, la mitad de los individuos es hipertenso o fumador. ¿Ser fumador es independiente de ser hipertenso en esta población?

Eventos independientes

Ejemplo. Hipertensos y fumadores

Sea

$$H = \text{Individuo hipertenso de la pob.}$$

$$F = \text{Individuo fumador de la pob.}$$

Si hubiera independencia, se tendría que $P(H \cap F) = P(H)P(F)$, pero como

$$P(H)P(F) = 0{,}2 \times 0{,}4 = 0{,}09$$

$$P(H \cap F) = P(H) + P(F) - P(H \cup F) = 0{,}2 + 0{,}45 - 0{,}5 = 0{,}15$$

los eventos no son independientes. En efecto, se nota que:

$$P(H|F) = \frac{0{,}15}{0{,}45} = 0{,}33$$

Mientras que $P(H) = 0,2$. Así, ser fumador aumenta la probabilidad de ser hipertenso.

Universidad del Rosario

Probabilidad total

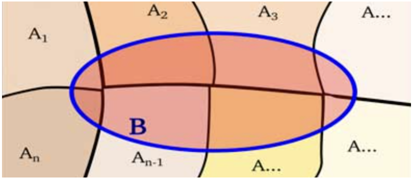
Definición

Al contar con una **partición** $A_1, ..., A_n$ sobre Ω:

1. Exhaustiva: $\cup A_i = \Omega$.
2. Mutuamente excluyente: $A_i \cap A_j = \emptyset$, $i \neq j$.

¿A qué es igual la probabilidad de cualquier evento B, inducida por la partición?

Figura 1: Evento B inducido por una partición

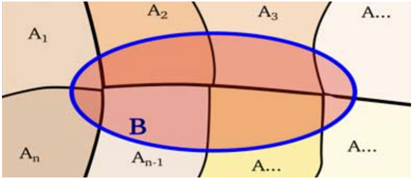


Probabilidad total

Definición

Suponga que $A_1, \ldots, A_n$ son las n causas posibles de que el evento B suceda (o las n ubicaciones posibles en las que el evento B puede ocurrir). El **teorema de probabilidad total** muestra la probabilidad de B como la suma de las probabilidades del evento B bajo todas las causas (o lugares) posibles $A_1, \ldots, A_n$.

Figura 2: Evento B inducido por una partición



Probabilidad total

Definición

Probabilidad total

Sea B un evento cualquiera y $A_1, ..., A_n$ una partición sobre Ω:

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Donde $P(A_i)$ se conocen como las probabilidades *apriori* (antes de que suceda B) de la partición. Se tiene además que $P(A_j|B)$ está dado por:

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

Son las probabilidades *aposteriori* (después de que suceda B) de la partición.

Universidad del Rosario

Ejemplo. Virus en dos zonas de la ciudad

Una ciudad está dividida en 2 zonas de salud: norte y sur, que atienden al 40 % y el 60 % de la población, respectivamente. Un estudio llevado a cabo en ambas zonas de salud revela que un 30 % de los individuos de la zona norte y un 40 % de los de la zona sur fueron infectados por covid. ¿Cuál es la probabilidad de que un individuo elegido al azar en esa ciudad hubiera tenido dicho virus?

Probabilidad total

Ejemplo

Ejemplo. Virus en dos zonas de la ciudad

En este caso, el evento B representa que el individuo tuvo el virus, mientras las partición dónde el evento puede suceder es: $A_1 =$ Individuo originario de la zona norte, y $A_2 =$ Individuo originario de la zona sur. Se pregunta $P(B)$:

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = 0{,}3 \times 0{,}4 + 0{,}4 \times 0{,}6 = 0{,}36$$

Con lo cual, hay un chance de 0.36 de que un individuo seleccionado al azar tenga el virus, ó, de manera equivalente, un 36 % de la población fue infectada por el virus.

Universidad del Rosario

Teorema de Bayes

Definición

En las probabilidades *aposteriori* presentadas, se sabe que

$$P(A_j \cap B) = P(B|A_j)P(A_j)$$

Y además

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Con lo cual se tiene el teorema de Bayes como sigue:

Teorema de Bayes

Sea B un evento cualquiera y $A_1, ..., A_n$ una partición sobre Ω. El teorema de Bayes indica que

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

Para $j = 1, ..., n$.

Ejemplo. Virus en dos zonas de la ciudad (cont.)

Si un habitante de la ciudad fue infectado por el coronavirus de estudio, ¿cuál es la probabilidad que provenga de la zona norte? ¿cuál es la probabilidad que provenga de la zona sur?

Ejemplo. Virus en dos zonas de la ciudad (cont.)

Se preguntan las probabilidades *aposteriori* de $P(A_1)$, con $P(A_1) = 0,40$, y de $P(A_2)$, con $P(A_2) = 0,60$. Es decir:

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^{2} P(B|A_i)P(A_i)} = \frac{0,3 \times 0,4}{0,36} = 0{,}33$$

$$P(A_2|B) = \frac{P(B|A_2)P(A_2)}{\sum_{i=1}^{2} P(B|A_i)P(A_i)} = \frac{0,4 \times 0,6}{0,36} = 0{,}66$$

Con lo cual, si un individuo tuvo el virus, es dos veces más probable que provenga de la zona sur que de la zona norte. Note además que la probabilidad de venir de la zona norte disminuye de un $40\,\% = P(A_1)$ a un $33\,\% = P(A_1|B)$, con el conocimiento de la pre existencia del virus.

Universidad del Rosario

Conclusiones

- La probabilidad de un evento puede cambiar al ser condicionada, múltiples ejemplos son observados en la vida diaria: chance diferencial de ocurrencia de un evento determinado entre razas, nivel educativo, hábitos (fumador, deportista, sedentario), entre otros.
- La probabilidad de un evento determinado se puede ver como la suma de probabilidades del evento condicionada bajo una partición determinada.

Universidad del Rosario

Contenido

Universidad del Rosario

Introducción

Recordemos algunas definiciones

1. Experimento aleatorio.
2. Espacio y espacio muestral.

Universidad del Rosario

Introducción

Al realizar un experimento aleatorio, Ω puede ser arbitrario ó tedioso de describir:

- Al lanzar una moneda $\Omega = \{C, S\}$ con C representando cara y S sello, ó quizá $\Omega = \{H, T\}$ con H representando cara y T sello.
- Al buscar el primer enfermo de un conjunto de personas $\Omega = \{(E), (S, E), (S, S, E), (S, S, S, E)...\}$ siendo E la persona es enferma y S, sana.

Universidad del Rosario

Introducción

Por otra parte, puede que no estemos interesados en los resultados del experimento aleatorio en si mismos, sino en ciertos **valores numéricos** a partir de ellos:

- Al lanzar dos monedas una vez $\Omega = \{(C, C), (C, S), (S, C), (S, S)\}$ puedo estar interesado en el número de caras obtenidas (0, 1 ó 2), más que en los $\omega \in \Omega$.
- Si nos interesan los ingresos de los servidores públicos de Colombia, al seleccionar uno de sus miembros, estamos interesados realmente en conocer su salario.
- ...

Universidad del Rosario

Introducción

Quizá sea buena idea hacer un **cambio de espacio**, en el cual sea menos difícil estudiar las propiedades de interés del experimento aleatorio:

- **Transportar** Ω a Ω', dónde Ω' es más fácil de estudiar.

Un espacio numérico que **estandarice** las propiedades de experimentos aleatorios similares.

Universidad del Rosario

Introducción

Ejemplo. Lanzamiento de una moneda una vez

Experimento aleatorio Lanzamiento de una moneda una vez.
Espacio muestral (de **salida**) $\Omega = \{C, S\}$.
Regla de asignación Número de caras observadas.
Espacio muestral (de **llegada**) $\Omega' = \{1, 0\}$.

Ejemplo. Lanzamiento de una moneda dos veces

Experimento aleatorio Lanzamiento de una monedas dos veces.
Espacio muestral (de **salida**) $\Omega = \{(C, C), (C, S), (S, C), (S, S)\}$.
Regla de asignación Número de caras observadas.
Espacio muestral (de **llegada**) $\Omega' = \{2, 1, 0\}$.

La idea es estudiar las características del experimento aleatorio ahora desde este nuevo espacio Ω', el cual es más sencillo de manipular.

Universidad del Rosario

Introducción

Ejemplo. Lanzamiento de una moneda una vez

Experimento aleatorio Lanzamiento de una moneda una vez.
Espacio muestral (de **salida**) $\Omega = \{C, S\}$.
Regla de asignación Número de caras observadas.
Espacio muestral (de **llegada**) $\Omega' = \{1, 0\}$.

Ejemplo. Lanzamiento de una moneda dos veces

Experimento aleatorio Lanzamiento de una monedas dos veces.
Espacio muestral (de **salida**) $\Omega = \{(C, C), (C, S), (S, C), (S, S)\}$.
Regla de asignación Número de caras observadas.
Espacio muestral (de **llegada**) $\Omega' = \{2, 1, 0\}$.

La idea es estudiar las características del experimento aleatorio ahora desde este nuevo espacio Ω', el cual es más sencillo de manipular.

Variable aleatoria

Definición

Variable aleatoria

Una regla (o también llamada *función*) X la cual está determinada por el resultado del experimento aleatorio, es decir, asigna a cada elemento $\omega \in \Omega$ un elemento $X(\omega) = x \in \Omega'$.

Ejemplo. Lanzamiento de una moneda dos veces

Experimento aleatorio Lanzamiento una moneda dos veces.
Espacio muestral $\Omega = \{(C, C), (C, S), (S, C), (S, S)\}$.
Variable aleatoria $X=$ *Número de caras observadas*.
Valores de la v.a

$$X((C, C)) = 2$$
$$X((C, S)) = 1$$
$$X((S, C)) = 1$$
$$X((S, S)) = 0$$

.

Variable aleatoria

Ejemplos

Algunos otros ejemplos de variables aleatorias (v.a.) son

1. $W =$ *Número de defectos observados en el proceso de fabricación*.
2. $Y =$ *Número de cestas realizadas en un partido de baloncesto*.
3. $Z =$ *Calificaciones de los estudiantes del curso de Estadística*.
4. $T =$ *Tiempo de espera en línea para la toma de un pedido*.

Ejemplo. Lanzamiento de una moneda dos veces

Volviendo al lanzamiento de una moneda dos veces, es claro que si la probabilidad de obtener cara es $p \in [0, 1]$, el vector de probabilidades en el espacio de salida Ω está dado por:

$$
\begin{aligned}
P(E_1) &= P(\{(C, C)\}) = p * p \\
P(E_2) &= P(\{(C, S)\}) = p * (1 - p) \\
P(E_3) &= P(\{(S, C)\}) = (1 - p) * p \\
P(E_4) &= P(\{(S, S)\}) = (1 - p) * (1 - p)
\end{aligned}
$$

Conformado por 4 eventos simples.

Ejemplo. Lanzamiento de una moneda dos veces

Por otra parte, el vector de probabilidades para el espacio de llegada Ω' está dado por:

$$
\begin{aligned}
P(X=2) &= P(\{\omega \in \Omega : X(\omega) = 2\}) = P(\{(C,C)\}) = p * p \\
P(X=1) &= P(\{\omega \in \Omega : X(\omega) = 1\}) = P(\{(C,S),(S,C)\}) \\
&= p * (1-p) + (1-p) * p = 2p * (1-p) \\
P(X=0) &= P(\{\omega \in \Omega : X(\omega) = 0\}) = P(\{(C,C)\}) = (1-p) * (1-p)
\end{aligned}
$$

Conformado por 3 eventos simples. En general asumimos que $\Omega' = \mathrm{R}$, con lo cual la variable X del ejemplo es una variable aleatoria **real**.

Contenido

Universidad del Rosario

Ejercicios pre-parcial

Responda Verdadero (V) o Falso (F) según corresponda:

- (F) Dos eventos simples cualesquiera son siempre independientes.
- (F) El teorema de Bayes muestra la probabilidad de B como la suma de las probabilidades del evento B bajo todas las causas (o lugares) posibles de una partición $A_1, ..., A_n$.
- (F) El teorema de probabilidad total muestra la probabilidad de todas las causas (o lugares) posibles de una partición $A_1, ..., A_n$ condicionado a la ocurrencia del evento B.

Universidad del Rosario

Ejercicios pre-parcial

Responda Verdadero (V) o Falso (F) según corresponda:

a. ( ) Un experimento no es aleatorio si se conoce su resultado de antemano.
b. ( ) Dos eventos simples diferentes son siempre mutuamente excluyentes.
c. ( ) Un evento compuesto se conforma mediante la unión de eventos simples.
d. ( ) El evento $(A \cap B)^c$ sucede cuándo pasa A^c, B^c ó ambos.
e. ( ) La estabilización de la regularidad estadística es igual a la frecuencia relativa.
f. ( ) La probabilidad siempre es mayor o igual que cero para cualquier evento.

Ejercicios pre-parcial

Se busca la primera persona con gripa de una población determinada de manera independiente, una vez encontrada, se detiene la búsqueda. ¿cuál es el espacio muestral asociado a este experimento aleatorio?. Si la probabilidad de tener gripa es p, ¿a qué es igual la probabilidad de un evento simple arbitrario E_i?