Probabilidad y Estadística 2019

Examen de entrenamiento 1 - Primer parcial

Problema 1 (25 %)

Escribid unas notas de clase que cubran los siguientes conceptos: concepto general de probabilidad condicionada, definición matemática de probabilidad condicionada, partición, teorema de la probabilidad total, sucesos independientes y probar el siguiente resultado:

Resultado: Si A y B son dos sucesos independientes, entonces también lo son A y $\overline{B}$ y $\overline{A}$ y $\overline{B}$.

Poned ejemplos de los conceptos desarrollados.

Las notas deben ser concisas y claras, como para un alumno que toma esta asignatura. La escritura, tanto en claridad como en corrección, se valorará especialmente.

Solución: Si A y B son independientes, por la condición necesaria y suficiente, $P(A \cap B) = P(A)P(B)$.

- El conjunto $A \cap \overline{B}$ se puede escribir como $A \cap \overline{B} = A - A \cap B$. Por las propiedades de la probabilidad y por la hipótesis de independencia, tenemos:

$$\begin{aligned} P(A \cap \overline{B}) =& P(A - A \cap B) = P(A) - P(A \cap B) \\ =& P(A) - P(A)P(B) = P(A) \cdot (1 - P(B)) = P(A)P(\overline{B}) \end{aligned}$$

  Esta igualdad prueba la independencia de A y $\overline{B}$.

- En esta última prueba haremos uso de la primera propiedad.

$$P(\overline{A}|\overline{B}) = 1 - P(A|\overline{B}) = 1 - P(A) = P(\overline{A})$$

Problema 2 (25 %)

(1) Sea X una variable aleatoria $G(p)$. Probad que la probabilidad de repetir el experimento k veces más antes de sacar el primer éxito no depende del número previo de veces que hayamos realizado el experimento.

Solución: Sea $h > 0$ y supongamos que se ha realizado el experimento k veces. El teorema establece que las probabilidades $P(X = k + h|X \geq k)$ y $P(X = k + h)$ son iguales. Esto es equivalente a decir que los resultados de los k primeros experimentos no influyen en los resultados que vienen a continuación. Si se ha realizado el experimento al menos k veces significa que X es mayor o igual que k; de lo contrario, ya se habría acabado el experimento. Tenemos que calcular cuánto vale $P(X = k + h|X \geq k)$.

$$\begin{aligned} P(X = k + h|X \geq k) =& \frac{P(\{X = k + h\} \cap \{X \geq k\})}{P(X \geq k)} = \frac{P(X = k + h)}{\sum_{i=k}^{\infty}(1 - p)^i \cdot p} \\ =& \frac{(1 - p)^{k+h} \cdot p}{\frac{(1-p)^k \cdot p}{p}} = (1 - p)^h \cdot p = P(X = h) \end{aligned}$$

(2) Tenemos dos instrumentos de medida de la distancia entre dos puntos. Las medidas de los dos instrumentos son variables aleatorias X_1, X_2 independientes con media común μ; se considera que μ es la distancia exacta. Por la experiencia pasada del uso de los instrumentos se acepta que las varianzas son conocidas e iguales a σ_1^2 y σ_2^2, respectivamente; estas varianzas no tienen por qué ser la misma. Dadas dos medidas de cada instrumento, la medida final que se toma es la media ponderada $\overline{\mu} = \alpha E(X_1) + (1-\alpha)E(X_2)$, donde α es un número real en $[0, 1]$. ¿Cómo se tiene que elegir α para que se minimice la varianza de la media ponderada?

Solución: Se define una variable X_3 como $X_3 = \alpha X_1 + (1-\alpha)X_2$, que es la media ponderada de ambas medidas. La media de X_3 es

$$E(X_3) = E(\alpha X_1 + (1-\alpha)X_2) = \alpha E(X_1) + (1-\alpha)E(X_2) = \alpha\mu + (1-\alpha)\mu = \mu$$

Se pide hallar la varianza, que se calcula como sigue (se tiene en cuenta la independencia de X_1 y X_2):

$$V(X_3) = V(\alpha X_1 + (1-\alpha)X_2) = \alpha^2 V(X_1) + (1-\alpha)^2 V(X_2) = \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2$$

Ahora se trata de minimizar esta expresión como función de α. Lo hacemos derivando e igualando a cero.

$$\begin{aligned}\frac{dV(X_3)}{d\alpha} =& 2\alpha\sigma_1^2 - 2(1-\alpha)\sigma_2^2 = \alpha(2\sigma_1^2 + 2\sigma_2^2) - 2\sigma_2^2;\\ & \alpha(2\sigma_1^2 + 2\sigma_2^2) - 2\sigma_2^2 = 0; \quad \alpha = \frac{\sigma_2^2}{\sigma_1^2 + 2\sigma_2^2}\end{aligned}$$

Hacemos la derivada segunda para confirmar que es un mínimo.

$$\frac{d^2V(X_3)}{d\alpha^2} = 2\sigma_1^2 + 2\sigma_2^2$$

Problema 3 (25 %)

Se tiene un experimento aleatorio en que solo hay dos sucesos de interés, A y su complementario. La probabilidad de A es constante en el tiempo. Se sabe que los resultados del experimento son independientes entre sí y se hacen n repeticiones del experimento. A continuación, se observan las siguientes variables:

a) X_1 es el número de veces que sale el suceso en A en las n repeticiones.

b) X_2 es el número de veces que se repite el experimento antes de obtener A.

c) X_3 es el número de veces que se repite el experimento antes de obtener la segunda A.

Se pide obtener razonadamente las funciones de masa de X_1, X_2 y X_3.
Solución:

(1) Obtener razonadamente las funciones de masa de X_1, X_2 y X_3.

a) Los resultados de los n experimentos de Bernouilli consecutivos e independientes se pueden identificar con las sucesiones de unos y ceros de longitud n, donde 1 representa éxito y 0 fracaso. Para una sucesión con k unos, existen $\binom{n}{k}$ posibilidades de colocar los unos en dicha sucesión. Tanto unos y ceros se colocan de modo independiente con probabilidad p y q, respectivamente. Por tanto,

$$P(X = k) = \binom{n}{k} p^k q^{n-k}$$

b) Dado que en la descripción de la variable pone número de veces *antes* repetir el experimento, X_2 toma valores en los números naturales, incluido el 0. Cada valor k de la variable está asociado con una sucesión de $k-1$ fracasos y un último éxito. Como hay independencia, tenemos:

$$P(X = k) = (1-p)^k \cdot p$$

c) X_3 mide el número de veces que se repite el experimento antes de obtener un segundo éxito. Se puede identificar los valores de esta variable con las sucesiones de longitud $k+1$ donde hay exactamente dos A y el resto son B, con una de las A en la última posición de la sucesión. Como los resultados son independientes entre sí en las probabilidades aparece p^2 y $(1-p)^{k-1}$. Sin embargo, falta colocar la primera A, que puede en cualquiera de las k primeras posiciones. Esa posición se puede elegir de $\binom{k}{1}$ maneras posible. Por tanto, la función de masa final es:

$$P(X_3 = k) = \binom{k}{1} \cdot p^2 \cdot (1-p)^{k-1}$$

Problema 4 (25 %)

Un test de diagnóstico tiene probabilidad 0,95 de dar resultado positivo cuando se le hace a una persona que sufre una cierta enfermedad y probabilidad 0,1 de un falso positivo cuando se le hace a una persona que no tiene esa enfermedad. Se estima que el 0,5 % de la población sufren esa enfermedad. Supongamos que el test se realiza a una persona de quien no se sabe nada con respecto a la enfermedad. Calcular las siguientes probabilidades:

(1) El test da positivo;

(2) Dado un resultado positivo, la persona sufre la enfermedad;

(3) Dado un resultado negativo, la persona no sufre la enfermedad;

(4) La persona ha recibido resultados erróneos del test.

Solución: Vamos a llamar P al suceso de que el test dé positivo, E que la persona sea portadora de la enfermedad, y F que el test dé resultados erróneos. Las probabilidades correspondientes son: $P(P|E) = 0{,}95$, $P(P|\overline{E}) = 0{,}1$ y $P(E) = 0{,}1$.

(1) Este es un problema que se puede resolver por Bayes. El espacio muestral es la población y está dividida en portadores de la enfermedad y no portadores. Por tanto, la probabilidad de positivo es:

$$P(P) = P(P|E) \cdot P(E) + P(P|\overline{E}) \cdot P(\overline{E}) = 0{,}95 \cdot 0{,}005 + 0{,}1 \cdot 0{,}995 = 0{,}10425$$

(2) Se pide la probabilidad de que sabiendo que el test ha dado un resultado positivo, la persona sufra la enfermedad. Esta es una probabilidad a posteriori, en concreto, $P(E|P)$.

$$\begin{aligned} P(E|P) =& \frac{P(P|E) \cdot P(E)}{P(P|E) \cdot P(E) + P(P|\overline{E}) \cdot P(\overline{E})} \\ =& \frac{0{,}95 \cdot 0{,}005}{0{,}95 \cdot 0{,}005 + 0{,}1 \cdot 0{,}995} = 0{,}0455 \end{aligned}$$

(3) Si el resultado es negativo, se pide la probabilidad de que la persona no sufra la enfermedad, esto es, $P(\overline{E}|\overline{P})$. Usaremos las propiedades de la probabilidad condicionada para este cálculo.

$$P(\overline{E}|\overline{P}) = \frac{P(\overline{E} \cap \overline{P})}{1 - P(P)} = \frac{P(\overline{P}|\overline{E}) \cdot P(\overline{P})}{1 - P(E)} = \frac{(1 - 0{,}1) \cdot 0{,}995}{1 - 0{,}10425} = 0{,}0997$$

(4) Una persona puede recibir resultados erróneos del test en dos circunstancias: uno, está sano y el test da positivo (falso positivo); dos, está enferma y el test da negativo (falso negativo). Entonces:

$$P(F) = P(P \cap \overline{E}) + P(\overline{P} \cap E) = P(P|\overline{E}) \cdot P(\overline{E}) + P(\overline{P}|E) \cdot P(E) = 0{,}09975$$