Estadística Bayesiana

Introducción

Dr Carlos López de Castilla Vásquez

Universidad Nacional Agraria La Molina

2022-1

Introducción

- La estadística Bayesiana le debe su nombre al trabajo pionero del reverendo Thomas Bayes titulado: *An Essay towards solving a Problem in the Doctrine of Chances* publicado en 1764.
- El artículo fue enviado a la Real Sociedad de Londres por Richard Price en 1763 quién escribió: *Yo ahora le mando un ensayo que he encontrado entre los papeles de nuestro fallecido amigo Thomas Bayes y el cual, en mi opinión, tiene un gran mérito y bien merece ser preservado* ...
- Aunque la obra de Thomas Bayes data ya de más de dos siglos la estadística Bayesiana es relativamente nueva y actualmente ostenta un gran desarrollo aunque no ajena también a grandes controversias.

Introducción

- La Estadística es el estudio de la *incertidumbre*. Una forma de lidiar con ella es pensar en las probabilidades.
- Algunos ejemplos: ¿cuál es la probabilidad de lanzar dos dados y obtener como suma cuatro? ¿cuál es la probabilidad que llueva mañana? ¿cuál es la probabilidad que el universo siga expandiéndose para siempre?
- Existen tres enfoques diferentes para definir probabilidades.
- La definición *clásica* se basa en la suposición de que cada resultado es igualmente probable.
- La definición *frecuentista* requiere que se tenga una secuencia hipotética de eventos infinitos.

Introducción

- La definición *Bayesiana* se basa en una perspectiva *personal*, donde la probabilidad es una medida del conocimiento que se tiene sobre un fenómeno de interés.
- Se trata de un enfoque *subjetivo*, pero que puede funcionar bien en una base estadística rigurosa, y que conduce a resultados mucho más intuitivos en comparación al enfoque frecuentista.
- En el ejemplo de la expansión del universo se podría considerar la teoría del multiverso, donde existe un número infinito de universos paralelos y preguntarse qué fracción tiene universos que se expanden para siempre. Dependiendo de lo que se sepa sobre el universo, se podrían obtener diferentes respuestas.

Teorema de Bayes

- Sean A y B dos eventos con $\Pr(B) > 0$, entonces:

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

- Si se tiene k eventos $A_1, \cdots, A_k$, los cuales constituyen una partición del espacio muestral Ω:

$$\Pr(A_i|B) = \frac{\Pr(B|A_i)\Pr(A_i)}{\Pr(B)} = \frac{\Pr(B|A_i)\Pr(A_i)}{\sum_{j=1}^{k}\Pr(B|A_j)\Pr(A_j)}$$

donde $\Pr(B)$ se le conoce como *probabilidad total*.

Teorema de Bayes

Ejemplo 1: Transmición de señales

- Un transmisor está enviando un mensaje mediante un código binario con señales de 0 y 1.
- Cada señal transmitida debe pasar por dos relevadores antes de llegar finalmente al receptor.
- En cada relevador, la probabilidad de que la señal enviada sea diferente de la señal recibida (inversión de señal) es de 0,10. Los relevadores funcionan independientemente uno de otro.
- El 60 % de todas las señales enviadas desde el emisor son 1.
- Si un 1 es recibido por el receptor ¿cuál es la probabilidad de que haya sido enviado realmente un 1?

Estadística Bayesiana

- El objetivo en la estadística Bayesiana es *representar* y *actualizar* la incertidumbre asociada a un parámetro θ.
- Un punto importante en la definición clásica de inferencia es que el parámetro θ es considerado como *constante*.
- La diferencia fundamental entre la teoría clásica y la Bayesiana está en que θ es considerado una *variable aleatoria*.
- La incertidumbre inicial sobre θ se representa usando una distribución inicial o *distribución a priori*.
- El proceso de actualización de la distribución a priori se realiza incorporando la información contenida en los datos, lo que permite hallar la *distribución posterior* para θ.

Estadística Bayesiana

- El teorema de Bayes, definido anteriormente usando eventos, puede presentarse en términos de funciones de probabilidad o densidad:
$$f(\theta|y) = \frac{f(y|\theta)\, f(\theta)}{f(y)}$$
- La distribución marginal para y se puede obtener de la siguiente manera:

$$f(y) = \begin{cases} \sum_\theta f(y|\theta)\, f(\theta) & \text{si } \theta \text{ es discreta} \\ \int_\theta f(y|\theta)\, f(\theta)\, d\theta & \text{si } \theta \text{ es continua} \end{cases}$$

y es llamada la *distribución predictiva a priori*.

Estadística Bayesiana

- La *distribución a priori* representa lo que es conocido de θ antes de observar los datos y se denota por $f(\theta)$.
- La *distribución posterior* representa lo que se conoce de θ después de observar los datos y se denota por $f(\theta|y)$.
- La función $f(y|\theta)$ es la *distribución muestral* e incorpora la información proporcionada por los datos.
- El núcleo de la distribución posterior es:

$$f(\theta|y) \propto f(y|\theta)\, f(\theta)$$

- En muchos casos la distribución muestral se reemplaza por la *función de verosimilitud* $L(\theta|y)$ en el teorema de Bayes.

Estadística Bayesiana

Ejemplo 2: Modelo binomial

- Suponga que se tiene una muestra:

$$y_1, \cdots, y_n \sim \mathcal{B}(\theta)$$

y que a priori $\theta \sim \mathcal{BE}(\alpha, \beta)$. Hallar la distribución posterior.

Ejemplo 3: Modelo de Poisson

- Suponga que se tiene una muestra:

$$y_1, \cdots, y_n \sim \mathcal{P}(\lambda)$$

y que a priori $\lambda \sim \mathcal{G}(\alpha, \beta)$. Hallar la distribución posterior.

Distribución predictiva posterior

- Luego de obtener la distribución posterior es posible realizar el proceso de predicción sobre una cantidad no observable $\tilde{y}$ que se define a partir de la distribución muestral.
- Algunas opciones son:

$$\tilde{y} = \widetilde{y}_i \quad \tilde{y} = \textstyle\sum_{i=1}^{m} \widetilde{y}_i \quad \tilde{y} = \frac{1}{m}\sum_{i=1}^{m} \widetilde{y}_i$$

- La *distribución predictiva posterior*:

$$f\left(\tilde{y}|y\right) = \begin{cases} \sum_\theta f\left(\tilde{y}|\theta\right) f\left(\theta|y\right) & \text{si } \theta \text{ es discreta} \\ \int_\theta f\left(\tilde{y}|\theta\right) f\left(\theta|y\right) d\theta & \text{si } \theta \text{ es continua} \end{cases}$$

Distribución predictiva posterior

Ejemplo 4: Distribución predictiva posterior modelo binomial

- Considere el modelo binomial del ejemplo 2. Hallar la distribución predictiva posterior para $\tilde{y} = \widetilde{y}_i$, un nuevo ensayo de Bernoulli.

Ejemplo 5: Distribución predictiva posterior modelo de Poisson

- Considere el modelo de Poisson del ejemplo 3. Hallar la distribución predictiva posterior para $\tilde{y} = \sum_{i=1}^{m} \widetilde{y}_i$, el número total de eventos de interés en una nueva muestra $\widetilde{y}_1, \cdots, \widetilde{y}_m$.