

Estadística Bayesiana

A. Mora, M. Ojeda

Departamento de Matemática Aplicada
Universidad de Málaga

Enfoque bayesiano vs enfoque frecuentista

- En el análisis de datos existe una distinción conceptual entre:
 - ▶ el contraste de hipótesis, y
 - ▶ la estimación con incertidumbre cuantificada.
- Existe también una segunda distinción conceptual entre
 - ▶ métodos frecuentistas, y
 - ▶ métodos bayesianos.

Existen razones para pasar de la “excesiva simplificación booleana” del contraste de hipótesis en función de la presencia o ausencia de efectos, a un enfoque acumulativo que vaya mejorando las estimaciones de variables y su intervalo de confianza.

Enfoque bayesiano vs enfoque frecuentista

- Todos los enfoques del análisis de datos están basados en modelos formales (matemáticos), con cierto número de parámetros.
- Los parámetros se pueden estudiar mediante contraste de hipótesis o por estimación con incertidumbre y, a su vez, mediante un enfoque frecuentista o bayesiano.

	Frecuentista	Bayesiano
Test de hipótesis	Valores p	Factor de Bayes
Estimación con incertidumbre	Máxima verosimilitud con intervalo de confianza	Distribución a posteriori con mayor intervalo de densidad

Valores p

- El típico test t requiere el cálculo de un valor p , de modo que si $p < 0,05$ la hipótesis nula es rechazada.
- Dado un dataset, se calcula el estadístico t . Para datos simples se tiene

$$t = \frac{\bar{y} - \mu_0}{s} \sqrt{N}$$

Lo que interesa es que t , en cierto modo, concentra información del dataset.

- El enfoque frecuentista determina la probabilidad de obtener este valor de t si la hipótesis nula es cierta.
- Cuando esta probabilidad p es suficientemente pequeña, la hipótesis nula se rechaza.
- El valor de p nos da la probabilidad de cometer un falso positivo (error de tipo I), y la cota elegida para p es para limitar los errores a ese 5%.

Estimación con intervalo de confianza

- En ocasiones nos puede interesar conocer la magnitud de la tendencia, no solo si la hipótesis nula es aceptable o no, y tener alguna idea sobre el grado de incertidumbre que esconden los datos.
- La mejor estimación de un parámetro en un modelo matemático se conoce como la *estimación de máxima verosimilitud* (MLE). Como ejemplo tenemos el estimador por mínimos cuadrados (LSE) que se usa en el análisis de varianza (ANOVA) o en regresión lineal.
- El intervalo de confianza nos mide hasta qué distancia del MLE del parámetro nos podemos alejar sin que la hipótesis nula sea rechazada.
- Los intervalos de confianza no forman una distribución de probabilidad. En particular, no nos dicen que valores del centro del intervalo sean más probables que los de los extremos.

Teorema de Bayes

- Probabilidad condicionada

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

- Probabilidad total

$$P(B) = \sum_{k=1}^n P(B/A_k) \cdot P(A_k)$$

- Teorema de Bayes

$$P(A_i/B) = \frac{P(B/A_i) \cdot P(A_i)}{P(B)} = \frac{P(B/A_i) \cdot P(A_i)}{\sum_{k=1}^n P(B/A_k) \cdot P(A_k)}$$

Estimación bayesiana I

(Intervalo de máxima densidad)

- El análisis bayesiano consiste en reubicar la credibilidad entre todas las posibilidades
- *Cuando eliminas lo imposible, lo que queda, por improbable que sea, debe ser verdad* (Sherlock Holmes)
- En análisis de datos, las posibilidades son valores de los parámetros de modelo descriptivo.
- Comenzamos asignando un grado de credibilidad a cada parámetro, sin considerar los datos. Esta la distribución *a priori*.
- La distribución obtenida después de procesar los datos es la distribución *a posteriori*.
- El nuevo reparto de credibilidad se hace mediante la fórmula de Bayes.
- La incertidumbre de la estimación viene determinada explícitamente por el intervalo de extensión de la distribución *a posteriori*.

Estimación bayesiana II

(Intervalo de máxima densidad)

- Se suele usar el intervalo de mayor densidad (HDI) al de 95%, que contiene los valores del parámetro de mayor probabilidad y que cubren el 95% de los valores más probables.
- Todo valor del parámetro dentro del HDI tiene mayor densidad de probabilidad (mayor credibilidad) que cualquier valor fuera del HDI.
- En este caso, las distribuciones obtenidas son probabilidades de los valores del parámetro, y no distribuciones muestrales de los datos. En particular, no tienen nada que ver con los valores p .

Estimación bayesiana

(Factores de Bayes)

- La hipótesis nula se expresa como una distribución a priori que pone toda la credibilidad en un pico (de densidad infinita) en el punto nulo del parámetro, y cero en los demás puntos.
- La hipótesis nula se compara con una distribución a priori alternativa (otro modelo).
- La inferencia bayesiana reubica la credibilidad sobre los modelos a priori y el parámetro.
- El grado por el cual el modelo se desplaza es el *factor de Bayes*.

$$BF_{null} = \frac{P(null/D)}{P(alt/D)} / \frac{P(null)}{P(alt)}$$

- La regla para decidir sobre la hipótesis nula se basa en el factor de Bayes, y no sobre las probabilidades posteriores.