Chequeo de Modelos y Análisis de Sensibilidad

Una vez que hemos llevado a cabo los dos pasos del Análisis Bayesiano:

- Construir un modelo de probabilidad
- Calcular (generalmente usando simulación) la distribución a posteriori de todos los estimandos

no debemos ignorar el paso final del chequeo del ajuste del modelo a los datos y a nuestro conocimiento del problema.

Nos referimos al **modelo** cuando hablamos de:

- la distribución de muestreo
- la distribución a priori
- cualquier estructura jerárquica
- otros aspectos tales como cuáles variables explicativas han sido incluidas en un modelo de regresión

Principios y métodos para el chequeo de un modelo

Se consideran tres formas mediante las cuales se puede utilizar la distribución a posteriori para chequear un modelo:

- Comparar la distribución a posteriori de los parámetros con un conocimiento importante del problema o con otros datos que no hayan sido incluidos en la distribución a priori o en la verosimilitud. **Ej**: Comparar la probabilidad a posteriori de θ_j= Probabilidad del jugador j de conectar un hit en un período de juego, con el rendimiento de los jugadores en períodos anteriores.
- Comparar la distribución predictiva a posteriori de observaciones futuras con el conocimiento del problema. **Ej**: Comparar las predicciones por estado de una elección presidencial con un conocimiento político importante sobre la preferencia de los candidatos en cada estado.

- Comparar la distribución predictiva a posteriori de observaciones futuras con los datos que ya han ocurrido. Esto implica que si el modelo es correcto los datos observados deben ser factibles bajo la distribución predictiva a posteriori. En este caso no se usa información adicional.

Comparación de los datos con la distribución predictiva a posteriori

Es muy útil hacer comparaciones gráficas de resúmenes de los datos con resúmenes de las simulaciones de la distribución predictiva a posteriori.

Sean : y^{rep}= una replicación de y (un dato que ha podido ser observado) y $\tilde{y}$= una observación futura observable. Por ejemplo, si el modelo tiene variables explicativas x, y y y^{rep} tienen iguales valores de x pero $\tilde{y}$ tiene posibles valores futuros $\tilde{x}$ de x.

Se trabaja con la distribución de y^{rep} dado los datos(conocimientos) actuales:

$$p(y^{rep}|y) = \int p(y^{rep}|\theta)p(\theta|y)d\theta$$

Pasos para medir las discrepancias enre los datos y las simulaciones de la distribución predictiva

- Se define una medida de discrepancia $T(y, \theta)$ que es una cantidad escalar que depende de los parámetros y los datos (equivalente a la estadística de prueba desde el punto de vista clásico).
- La falta de ajuste de los datos con respecto a la distribución predictiva se mide con el p-valor o la probabilidad de la cola.
  - Desde un punto de vista clásico:
    p-valor= $Pr(T(y^{rep}) \geq T(y)|\theta)$, donde la probabilidad se calcula sobre la distribución de y^{rep} con θ fijo.

- Desde un punto de vista Bayesiano:

  p-valor predictivo a posteriori= $Pr(T(y^{rep}, \theta) \geq T(y, \theta)|y)$

  donde la probabilidad se toma sobre la distribución a posteriori de θ y la distribución a posteriori predictiva de y^{rep}.

  Esta es la probabilidad de que las réplicas sean más extremas que los datos.

Una medida general de discrepancia es la discrepancia χ^2 o prueba de bondad de ajuste es:

$$\mathrm{d}iscrepancia\ \ \chi^2 : \ T(y,\theta) = \sum_i \frac{(y_i - E(y_i|\theta))^2}{var(y_i|\theta)}$$

La distribución χ^2 de referencia está basada en la aproximación para nuestras grandes de la distribución a posteriori.

Interpretación de los p-valores predictivos a posteriori

- Un modelo es dudoso si la probabilidad de la cola para alguna cantidad de prueba de interés está cercano a 0 o a 1 (menor que 0.01 o mayor 0.99). Los p-valores no deben ser interpretados como $Pr(Modelo\ es\ verdadero|Datos)$.
- Los p-valores no deben ser interpretados como evidencia numérica. Es decir, un p-valor de 0.00001 no es más fuerte en la práctica que 0.001. En ambos casos el aspecto de los datos medido para la cantidad de prueba es inconsistente con el modelo.
- The objetivo más importante no es responder a la pregunta: "¿provienen los datos del modelo propuesto?", si no, cuantificar las discrepancias entre los datos y el modelo y determinar si estas discrepancias provienen del azar bajo las suposiciones del mismo modelo.

Relación de las pruebas clásicas con las pruebas Bayesianas

Las pruebas Bayesianas predictivas a posteriori son generalizaciones de las pruebas clásicas ya que se promedia sobre la distribución a posteriori del vector de parámetros desconocido en lugar de sobre algún valor fijo $\hat{\theta}$. Estas pruebas no dependen de una cantidad pivotal predeterminada ni de resultados asintóticos y por lo tanto pueden aplicarse a cualquier modelo de probabilidad.

Pero, cuidado! Estas pruebas **no son** automáticas: la selección de una cantidad de prueba y de una distribución predictiva apropiada necesita considerar el tipo de inferencia que se requiere para el problema en cuestión.

Análisis de sensibilidad

Se considera la incertidumbre en las inferencias a posteriori debido a la existencia de modelos alternativos razonables.

Se pueden considerar otros modelos que difieran en la especificación de la distribución a priori, la verosimilitud o ambos.

El método básico de análisis de sensibilidad es ajustar varios modelos de probabilidad al mismo problema.

Ejemplos:

- Reemplazar distribuciones a priori impropias con un conocimiento a priori importante sobre el problema.
- Usar modelos robustos que aseguren poca influencia de observaciones inusuales, como por ejemplo, usar una distribución t-Student en lugar de la normal

Comparación de un conjunto discreto de modelos usando factores de Bayes

Factor de Bayes: Es el cociente de la verosimilitud marginal de un modelo con respecto a la verosimilitud marginal de otro modelo.

Sean H_1 y H_2 dos modelos que compiten; el radio de sus probabilidades a posteriori es:

$$\frac{p(H_2|y)}{p(H_1|y)} = \frac{p(H_2)}{p(H_1)} x\mathrm{F}actordeBayes(H_2;H_1)$$

donde

$$\mathrm{F}actordeBayes(H_2;H_1) = \frac{p(y|H_2)}{p(y|H_1)} = \frac{\int p(\theta_2|H_2)p(y|\theta_2,H_2)d\theta_2}{\int p(\theta_1|H_1)p(y|\theta_1,H_1)d\theta_1}$$

El factor de Bayes está definido cuando la densidad marginal de y sobre cada modelo es propia.

Un ejemplo donde los factores de Bayes son de utilidad:
Ejemplo de genética de la sección 1.4 del GCSR:

Modelos que compiten: H_1: La mujer está afectada; y H_2: La mujer no está afectada; esto es: $\theta = 1$ y $\theta = 0$.

La razón de probabilidad a priori es $p(H_2)/p(H_1) = 1$.

El factor de Bayes de los datos en los que la mujer tiene dos hijos afectados es: $p(y|H_2)/p(y|H_1) = 0{,}25/1{,}0$.

El radio a posteriori es $p(H_2|y)/p(H_1|y) = 0.25$.

En este caso no hay modelos intermedios entre los dos modelos a comparar y cada $p(y|H_i)$ es propia.

Ejercicio: Analizar el caso de los factores de Bayes para el ejemplo 5.5 del GCSR (pag. 176) y el chequeo del modelo para este problema (sección 6.8).