Modelos bayesianos - 9 de enero de 2024

Ejercicio 1.a (1 punto) En un estudio sobre preferencias de viaje se ha investigado la relación entre el tipo de destino (urbano, rural), el transporte (coche, tren, avión), la frecuencia de viajes realizados al año (mucha, poca) y la satisfacción con las experiencias del viaje (alta, baja). En dicho análisis se han llegado a estas conclusiones:
Pr(Transporte = avión) = 0.5,
Pr(Transporte = coche) = 0.3,
Pr(Transporte = tren) = 0.2,
Pr(Frecuencia = mucha) = 0.3,
Pr(Frecuencia = poca) = 0.7,
Pr(Destino = urbano | Transporte = avión, Frecuencia = mucha) = 0.8,
Pr(Destino = urbano | Transporte = coche, Frecuencia = mucha) = 0.6,
Pr(Destino = urbano | Transporte = tren, Frecuencia = mucha) = 0.4,
Pr(Destino = urbano | Transporte = avión, Frecuencia = poca) = 0.4,
Pr(Destino = urbano | Transporte = coche, Frecuencia = poca) = 0.2,
Pr(Destino = urbano | Transporte = tren, Frecuencia = poca) = 0.1,
Pr(Satisfacción = baja | Destino = urbano) = 0.2,
Pr(Satisfacción = baja | Destino = rural) = 0.4.

Justifique adecuadamente las respuestas.

1. ¿Cuál es la probabilidad de que el destino sea rural? (a) 0.606 (b) 0.5 (c) 0.394.
2. ¿Cuál es la probabilidad de que el destino sea urbano si se viaja en avión? (a) 0.52 (b) 0.5 (c) 0.48.

Ejercicio 1.b (1 punto)

> «Fue seguramente una candidez, quizá necesaria, de la Ciencia naciente el de imaginarse que podría observar los fenómenos en sí mismos, tal como se desarrollarían fuera de nosotros mismos. Instintivamente, los físicos y los naturalistas operaron al principio como si su mirada cayera

desde lo alto sobre un Mundo en el que su conciencia pudiera penetrar sin experimentarlo en sí mismos, sin modificarlo con su propia observación. Hoy empiezan a darse cuenta de que sus observaciones, aun las más objetivas, están todas ellas impregnadas de convenciones apriorísticas, así como de formas o costumbres de pensar desarrolladas a lo largo del proceso histórico de la Investigación. Llegados al extremo de su análisis, ya no están seguros de si la estructura conseguida es la esencia misma de la Materia que estudian o el reflejo de su propio pensamiento. [...] El objeto y el sujeto se mezclan y se transforman mutuamente en el acto del conocimiento.» P. Teilhard de Chardin, *El fenómeno humano*, 1955, pág. 44, ed. Taurus.

Reflexione sobre el texto anterior de acuerdo a los principios bayesianos.

Ejercicio 2. (2 puntos) El rey Basilio ha encargado al príncipe Segismundo un estudio sobre la proporción de personas del reino que creen en la profecía de que el fin del mundo se acerca. Tras entrevistar a 120 personas, resulta que 74 aseguran que el fin del mundo se acerca, mientras que el resto de encuestados opina lo contrario. Segismundo redacta un informe donde detalla el estimador más verosímil y el intervalo de confianza al 95 %. Rosaura, una extranjera recién llegada al reino y que no conoce nada del mismo, lee ese informe y elabora otro alternativo presuponiendo estocasticidad bayesiana en las magnitudes analizadas. Responda razonadamente a las siguientes preguntas.

1. (0.5 puntos) Reproduzca numéricamente el informe de Segismundo. ¿Cómo se interpreta el intervalo de confianza?
2. (1 punto) Describa la metodología detallada en el informe de Rosaura. ¿Cuál es la máxima estimación a posteriori?
3. (0.5 puntos) Comparando ambos enfoques, ¿qué ventajas e inconvenientes presentan estos informes?

Ejercicio 3. (2 puntos) Sea $J(\theta_1 \mid \theta_0)$ la distribución probabilística definida como

$$J(\theta_1 \mid \theta_0) \sim \mathit{Uniforme}(\theta_0 - \delta, \theta_0 + \delta), \quad \delta > 0.$$

1. (1 punto) Demuestre que esta distribución cumple que $J(\theta_1 \mid \theta_0) = J(\theta_0 \mid \theta_1)$, i.e., la probabilidad de obtener $\theta_1 = \theta^a$ dado que $\theta_0 = \theta^b$ es igual a la probabilidad de obtener $\theta_1 = \theta^b$ dado que $\theta_0 = \theta^a$.
2. (0.2 puntos) ¿Cómo se denomina la propiedad anterior, $J(\theta_1 \mid \theta_0) = J(\theta_0 \mid \theta_1)$?
3. (0.8 puntos) De los siguientes procedimientos, –teorema de Bayes, estimador máximo verosímil, intervalo de confianza, intervalo de credibilidad, máximo a posteriori, simulación de Montecarlo, Metropolis, Metropolis-Hastings, Gibbs, Montecarlo hamiltoniano, digrafo acíclico–, ¿en cuál(es) de ellos aplicaría, o no, esta función? ¿con qué finalidad? ¿por qué? Razone adecuadamente las respuestas.

Ejercicio 4. (2 puntos) Sea $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)$ un conjunto de variables aleatorias tales que siguen la misma distribución, $\beta_0, \beta_1, \ldots, \beta_p \sim \mathit{Normal}(0,1)$ y son independientes. Se define la variable aleatoria Y de la siguiente forma

$$p(y_i \mid \boldsymbol{\beta}) \propto \exp\left(-\frac{1}{18}(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta})^2\right),$$

con $\boldsymbol{x}_i = (x_{i,0}, x_{i,1}, \ldots, x_{i,p})$ vectores de datos fijos para $i = 1, \ldots, n$ e y_i independientes entre sí. Justifique adecuadamente las respuestas.

1. Explicite la distribución conjunta de $\boldsymbol{\beta}$.
2. Determine una expresión proporcional a la distribución de

$$p(y_1, y_2, \ldots, y_n \mid \boldsymbol{\beta}).$$

3. Determine una expresión proporcional a $p(\boldsymbol{\beta} \mid y_1, y_2, \ldots, y_n)$.

4. El estimador máximo a posteriori de $\boldsymbol{\beta}$, ¿sirve como solución $\hat{\boldsymbol{\beta}}_{\text{ridge}}$ a una regresión de tipo ridge?

$$\hat{\boldsymbol{\beta}}_{\text{ridge}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i^T \boldsymbol{\beta})^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}.$$

Ejercicio 5. (2 puntos) Considere el problema consistente en estudiar la probabilidad de aprobar de un determinado estudiante teniendo en cuenta las variables «Tiempo dedicado a la asignatura por parte del estudiante» (T), «Asistencia a clase por parte del estudiante» (A), «Claridad de la clase expositiva» (E), «Calidad de los materiales docentes» (M) y «Resultado de la evaluación del estudiante» (R).

1. (0.2 puntos) Proponga, en primer lugar, para cada una de las variables, un conjunto razonable de valores que pueden tomar procurando, para mejorar la calidad del modelo, que no todas ellas sean dicotómicas.
2. (0.25 puntos) Establezca las relaciones causales que le resulten más adecuadas entre las variables a través de un grafo acíclico dirigido y explique brevemente por qué considera ésas y no otras.
3. (0.25 puntos) Proponga unas tablas de probabilidades condicionadas coherentes con el grafo del apartado anterior.
4. (0.4 puntos) Considere el conjunto de variables $X = \{T, A\}$. ¿Están d-separadas las variables R y M dado X en su red bayesiana? ¿Depende esa d-separación de los valores numéricos que haya asignado en las tablas?
5. (0.4 puntos) Determine, a partir de las tablas de probabilidades condicionadas de la red, las distribuciones marginales a priori de cada una de las variables involucradas.
6. (0.5 puntos) ¿Qué ocurriría con las distribuciones marginales de la variable R para una cierta asignatura a cuyo estudio sabemos que el alumno en cuestión ha dedicado el máximo de tiempo posible? ¿Como se reflejaría eso en las probabilidades de la red? ¿Y si es sabido que el alumno ha sacado la máxima nota posible, cómo afectaría a la distribución de probabilidad de la variable A?