Universidad de Oviedo

Modelos bayesianos. Junio de 2024

Ejercicio 1 (3 puntos). 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 la satisfacción sea baja? (a) 0.3212 (b) 0.19 (c) 0.7.
2. Demuestre que las variables Satisfacción y Transporte son condicionalmente independientes dado el Destino si y solo si Pr(Satisfacción | Transporte, Destino) = Pr(Satisfacción | Destino).
3. Supóngase que Pr(Satisfacción = baja) = 0.3212 y que Pr(Transporte = tren, Destino = rural, Satisfacción = baja) = 0.0648. ¿Cuál es la probabilidad de que el transporte sea el tren si la satisfacción es baja? (a) 0.2254 (b) 0.362 (c) 0.638.

Ejercicio 2 (1 punto). Reflexione sobre el siguiente texto y proponga argumentos a favor o en contra sobre las conclusiones que obtiene.

> «Dícese que los griegos distinguían tres (cuando menos) tipos de conocimiento:
>
> - Doxa: o aquello que conocemos porque nos lo han contado, sea en Twitter o en arXiv.
> - Gnosis: o aquello que conocemos por la experiencia personal, a través de los sentidos o, supongo que hoy en día, también a través de instrumentos de medida diversos.
> - Episteme: o aquello que decimos saber porque hemos razonado y tenemos ciertas garantías de su veracidad.
>
> Así planteados, son tres patas de un mismo taburete, tres monedas en el bolsillo, un conjunto, en definitiva, de tres elementos.
>
> »Pero, ¿de dónde puede venir la episteme si no es a través de una síntesis de las otras dos? No es difícil intuir una estructura en la tríada donde la episteme, la posteriori, es producto de la adecuada combinación de la doxa, actuando como priori, y de la gnosis, como los datos observados.» Carlos Gil Bellosta, *Doxa, episteme y gnosis: una reinterpretación bayesiana*, 2024, https://www.datanalytics.com/.

Ejercicio 3 (2 puntos). Sea $\theta \sim Uniforme(0,1)$, $Y \sim Binomial(57,\theta)$, y la evidencia obtenida es $y = 19$. Se desea emplear el algoritmo de Metropolis para simular la distribución a posteriori. Justifique adecuadamente las respuestas.

1. (0.5 puntos) Explique cómo se interpreta la razón de aceptación del algoritmo Metropolis.
2. (0.5 puntos) Demuestre en qué casos el ratio de aceptación del algoritmo Metropolis-Hastings es igual al del algoritmo Metropolis.

3. (1 punto) Sea s una iteración arbitraria del algoritmo Metropolis. Supongamos que $\theta^{(s-1)} = 0.8$ y el valor candidato en la iteración s es $\theta^{(*)} = 0.4$. Determine el valor de $\theta^{(s)}$ y el razonamiento seguido para obtenerlo.

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, $\beta_0, \beta_1, \ldots, \beta_p \sim Normal(0,1)$, y son independientes. La variable $\tau \sim Gamma(1/2, 1/2)$ es independiente de $\boldsymbol{\beta}$. Se define la variable aleatoria $\boldsymbol{y}$ como $p(y_i \mid \boldsymbol{\beta}, \tau) \propto \exp\left(-\frac{\tau}{2}(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta})^2\right)$, para y_i independientes entre sí, $i = 1, \ldots, n$. Justifique adecuadamente las respuestas.

1. ¿Cuál es la distribución conjunta de $\boldsymbol{\beta}$? ¿Cuánto vale su media y su matriz de covarianzas?
2. ¿Qué distribución sigue $p(\boldsymbol{y} \mid \boldsymbol{\beta}, \tau)$?
3. Demuestre que $p(\boldsymbol{\beta} \mid \boldsymbol{y}, \tau) \propto p(\boldsymbol{y} \mid \boldsymbol{\beta}, \tau) p(\boldsymbol{\beta})$.
4. ¿Cuál es la distribución de $p(\boldsymbol{\beta} \mid \boldsymbol{y}, \tau)$?

Notación y resultados previos: $\boldsymbol{x}_i = (x_{i,0}, x_{i,1}, \ldots, x_{i,p})$ vectores de datos fijos, $\boldsymbol{X} = (\boldsymbol{x}_i)_{i=1,\ldots,n}$, $\boldsymbol{y} = (y_1, \ldots, y_n)$, $(\boldsymbol{\beta} - \boldsymbol{0})^T \boldsymbol{I}^{-1} (\boldsymbol{\beta} - \boldsymbol{0}) = \boldsymbol{\beta}^T \boldsymbol{\beta} = \sum_{j=0}^{p} \beta_j^2$ con $\boldsymbol{I}$ la matriz identidad, $SSR(\boldsymbol{\beta}) = \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i^T \boldsymbol{\beta})^2 = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^T (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) = \boldsymbol{y}^T \boldsymbol{y} - 2\boldsymbol{\beta}^T \boldsymbol{X}^T \boldsymbol{y} + \boldsymbol{\beta}^T \boldsymbol{X}^T \boldsymbol{X} \boldsymbol{\beta}$, y $\boldsymbol{\theta} \sim Normal\ Multivariante(\boldsymbol{\mu}, \sqrt{\Lambda})$ si y solo si

$$
\begin{aligned}
p(\boldsymbol{\theta}) &= \frac{1}{(2\pi)^{r/2} \mid \Lambda \mid^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\mu})^T \Lambda^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu})\right) \\
&\propto \exp\left(-\frac{1}{2}\boldsymbol{\theta}^T \Lambda^{-1} \boldsymbol{\theta} + \boldsymbol{\theta}^T \Lambda^{-1} \boldsymbol{\mu}\right).
\end{aligned}
$$

Algunas distribuciones normales multivariantes de interés:

$$
\begin{cases}
Normal\ Multivariante\,(\boldsymbol{0}, \boldsymbol{I}), \\
Normal\ Multivariante\left(\boldsymbol{X}\boldsymbol{\beta}, \frac{1}{\sqrt{\tau}}\boldsymbol{I}\right), \\
Normal\ Multivariante\left(\frac{1}{\boldsymbol{X}^T\boldsymbol{X}\tau + \boldsymbol{I}} \boldsymbol{X}^T \boldsymbol{y} \tau, \sqrt{\frac{1}{\boldsymbol{X}^T\boldsymbol{X}\tau + \boldsymbol{I}}}\boldsymbol{I}\right).
\end{cases}
$$

Ejercicio 5 (2 puntos). Se quiere modelar una red bayesiana para decidir si un correo es basura o no. Se decide, por simpleza y operatividad, hacerlo a través de una red bayesiana ingenua (*Naïve Bayes structure*). En la red se incluyen las variables dicotómicas S, G, U y E que toman los valores «Sí» o «No» en función de si el correo se clasifica como basura (*spam*) o no, de si el mensaje contiene la palabra «Gratis» o no, de si texto incluye la expresión «Urgente» o no y de si el correo procede de un remitente extraño o no, respectivamente.

Los datos de entrenamiento del modelo son 1 000 correos de un determinado usuario, de los cuales 200 resultaron ser *spam*. De estos 200, 100 incluían la palabra «Gratis», 30 contenían la expresión «Urgente» y 180 procedían de remitentes extraños. De los mensajes que no fueron *spam*, en el 5 % aparecía la palabra «Gratis», en el 1 % «Urgente» y el 20 % provenían de un remitente extraño.

Con la información proporcionada en el enunciado responda a las siguientes preguntas.

1. (0.3 puntos) Represente, mediante un grafo acíclico dirigido, la red bayesiana ingenua descrita.
2. (0.3 puntos) Proporcione las distribuciones de probabilidad de los cuatro sucesos involucrados mediante las tablas de probabilidad condicionadas de la red.
3. (0.3 puntos) Determine la cobertura de Markov de cada variable de la red.
4. (0.3 puntos) ¿Están d-separadas, dos a dos, las variables G, U y E dada S? ¿Son codicionalmente independientes, dos a dos, las variables G, U y E dada S?
5. (0.8 puntos) Determine la probabilidad de que el correo «Reciba gratis y urgente un paquete de regalo» remitido desde una dirección extraña sea clasificado como *spam*.