Guia Examen Baidya

1

1. Se asume que los datos bivariados (4, 10), (-1, 3), (0, 2) surgen del modelo $y_i = b|x_i - 3| + e_i$, donde b es una constante y e_i son variables aleatorias independientes.

   (a) ¿Qué suposiciones son necesarias en e_i para que tenga sentido hacer un ajuste por mínimos cuadrados de una curva $y = b|x - 3|$ a los datos?

   (b) Dados los datos anteriores, determinar la estimación de mínimos cuadrados para b.

Respuesta

Problem 11. (15) **(a)** We assume the random error terms e_i are independent, have mean 0 and all have the same variance (homoscedastic).

(b)

$$
\begin{aligned}
E(b) &= \text{sum of the squared errors} \\
&= \sum (y_i - b|x_i - 3|)^2 \\
&= (10 - b)^2 + (3 - 4b)^2 + (2 - 3b)^2
\end{aligned}
$$

The least squares fit is found by setting the derivative (with respect to b) to 0,

$$\frac{d\,E(b)}{db} = -2(10 - b) - 8(3 - 4b) - 6(2 - 3b) = 52b - 56 = 0.$$

Therefore the least squares estimate of b is $\boxed{\hat{b} = \frac{56}{52} = \frac{14}{13}}$.

2

2. Una moneda se lanza 100 veces y resulta en cara 62 veces. ¿Cuál es la estimación de máxima verosimilitud para la probabilidad de resultar cara?

Respuesta

6. (a) The likelihood function is

$$p(\text{data}|\theta) = \binom{100}{62}\theta^{62}(1-\theta)^{38} = c\theta^{62}(1-\theta)^{38}.$$

To find the MLE we find the derivative of the log-likelihood and set it to 0.

$$\ln(p(\text{data}|\theta)) = \ln(c) + 62\ln(\theta) + 38\ln(1-\theta).$$

$$\frac{d\,\ln(p(\text{data}|\theta))}{d\theta} = \frac{62}{\theta} - \frac{38}{1-\theta} = 0.$$

The algebra leads to the MLE $\boxed{\theta = 62/100}$.

(b) The computation is identical to part (a). The likelihood function is

$$p(\text{data}|\theta) = \binom{n}{k}\theta^{k}(1-\theta)^{n-k} = c\theta^{k}(1-\theta)^{n-k}.$$

To find the MLE we set the derivative of the log-likelihood and set it to 0.

$$\ln(p(\text{data}|\theta)) = \ln(c) + k\ln(\theta) + (n-k)\ln(1-\theta).$$

$$\frac{d\,\ln(p(\text{data}|\theta))}{d\theta} = \frac{k}{\theta} - \frac{n-k}{1-\theta} = 0.$$

The algebra leads to the MLE $\boxed{\theta = k/n}$.

3

3. Digamos que en un curso, la probabilidad de que cada estudiante obtenga una nota determinada es:

$$P(A) = 1/2, \quad P(B) = \mu, \quad P(C) = 2\mu, \quad P(D) = 1/2 - 3\mu$$

(a) A continuación, observamos algunos datos y sean a, b, c y d el número de estudiantes que obtuvieron A, B, C y D, respectivamente. ¿Cuál es el estimador de máxima verosimilitud para μ, dado a, b, c y d?

(b) Ahora supongamos que observamos algunos datos nuevos, pero no observamos a ó b. En cambio, observamos h, que es el número de estudiantes que obtuvieron entre A ó B. Así que no conocemos a ó b, pero sabemos que $h = a + b$ y todavía observamos c y d como antes. Ahora queremos usar el algoritmo EM para obtener una estimación de μ.

i. Paso de la esperanza: Dado $\hat{\mu}$, una estimación actual de μ, ¿cuáles son los valores esperados de a y b?

ii. Paso de la maximización: Dado $\hat{a}$ y $\hat{b}$, los valores esperados de a y b, ¿cual es estimador de máxima verosimilitud de μ?

Respuesta

P(A) = ½ P(B) = μ P(C) = 2μ P(D) = ½-3μ

P(*a*,*b*,*c*,*d* | μ) = $K(\frac{1}{2})^a(\mu)^b(2\mu)^c(\frac{1}{2}-3\mu)^d$

log P(*a*,*b*,*c*,*d* | μ) = log K + *a*log ½ + *b*log μ + *c*log 2μ + *d*log (½-3μ)

$$\text{FOR MAX LIKE } \mu, \text{ SET } \frac{\partial \text{LogP}}{\partial \mu} = 0$$

$$\frac{\partial \text{LogP}}{\partial \mu} = \frac{b}{\mu} + \frac{2c}{2\mu} - \frac{3d}{1/2 - 3\mu} = 0$$

$$\text{Gives max like } \mu = \frac{b+c}{6(b+c+d)}$$

So if class got

A	B	C	D
14	6	9	10

$$\text{Max like } \mu = \frac{1}{10}$$

b)

We can answer this question circularly:

EXPECTATION

If we know the value of μ we could compute the expected value of *a* and *b*

Since the ratio a:b should be the same as the ratio ½ : μ

$$a = \frac{\frac{1}{2}}{\frac{1}{2} + \mu} h \qquad b = \frac{\mu}{\frac{1}{2} + \mu} h$$

MAXIMIZATION

If we know the expected values of *a* and *b* we could compute the maximum likelihood value of μ

$$\mu = \frac{b+c}{6(b+c+d)}$$

We begin with a guess for μ

We iterate between EXPECTATION and MAXIMALIZATION to improve our estimates of μ and *a* and *b*.

Define $\mu^{(t)}$ the estimate of μ on the t'th iteration

$b^{(t)}$ the estimate of *b* on t'th iteration

$$\mu^{(0)} = \text{ initial guess}$$

$$b^{(t)} = \frac{\mu^{(t)} h}{\frac{1}{2} + \mu^{(t)}} = \mathrm{E}\left[b \mid \mu^{(t)}\right]$$

E-step

$$\mu^{(t+1)} = \frac{b^{(t)} + c}{6\left(b^{(t)} + c + d\right)}$$

$$= \text{ max like est. of } \mu \text{ given } b^{(t)}$$

M-step

Continue iterating until converged.
Good news: Converging to local optimum is assured.
Bad news: I said "local" optimum.

4

4. Supongamos que nos gustaría tomar una muestra de la siguiente distribución conjunta, π de X y Y usando muestreo de Gibbs:

$$P(X=0,Y=0)=1/6,\quad P(X=0,Y=1)=1/8,\quad P(X=1,Y=0)=1/4$$

$$P(X=1,Y=1)=1/4,\quad P(X=2,Y=0)=1/12,\quad P(X=2,Y=1)=1/8$$

Calcular la distribuciones condicional

$$\pi_1(x|y)=P(X=x|Y=y)\ \text{ y }\ \pi_2(y|x)=P(Y=y|X=x)\ \text{ para }\ x\in\{0,1,2\},\quad y\in\{0,1\}$$

Respuesta

3. Suppose we would like to sample the following joint distribution π of X and Y using Gibbs Sampling.

	$X=0$	$X=1$	$X=2$
$Y=0$	$1/6$	$1/4$	$1/12$
$Y=1$	$1/8$	$1/4$	$1/8$

a) Compute the full conditional distributions

$$\begin{aligned}\pi_1(x|y) &= \mathbb{P}(X=x|Y=y) \quad \text{and} \\ \pi_2(y|x) &= \mathbb{P}(Y=y|X=x) \quad \text{for } x\in\{0,1,2\} \text{ and } y\in\{0,1\}.\end{aligned}$$

The first conditional distribution is given by

$$\begin{aligned}\pi_1(0|0) &= \frac{1/6}{1/6+1/4+1/12} = 1/3 & \pi_1(1|0) &= \frac{1/4}{1/6+1/4+1/12} = 1/2 \\ \pi_1(2|0) &= \frac{1/12}{1/6+1/4+1/12} = 1/6 & \pi_1(0|1) &= \frac{1/8}{1/8+1/4+1/8} = 1/4 \\ \pi_1(1|1) &= \frac{1/4}{1/8+1/4+1/8} = 1/2 & \pi_1(2|1) &= \frac{1/8}{1/8+1/4+1/8} = 1/4\end{aligned}$$

The second conditional distribution is given by

$$\begin{aligned}\pi_2(0|0) &= \frac{1/6}{1/6+1/8} = 4/7 & \pi_2(1|0) &= \frac{1/8}{1/6+1/8} = 3/7 \\ \pi_2(0|1) &= \frac{1/4}{1/4+1/4} = 1/2 & \pi_2(1|1) &= \frac{1/4}{1/4+1/4} = 1/2 \\ \pi_2(0|2) &= \frac{1/12}{1/12+1/8} = 2/5 & \pi_2(1|2) &= \frac{1/8}{1/12+1/8} = 3/5\end{aligned}$$

b) The full conditional distributions from a) can be used to define a Gibbs sampler chain $(X_n, Y_n)_{n\geq 0}$ whose stationary distribution is π. Suppose the states of the Markov chain $(X_n, Y_n)_{n\geq 0}$ are labelled

$$1 = (0,0),\ 2 = (1,0),\ 3 = (2,0),\ 4 = (0,1),\ 5 = (1,1),\ 6 = (2,1).$$

Compute the transition matrix for the update of the first component and the transition matrix for the update of the second component.
The update of the first component is according to the transition matrix

$$P_1 = \begin{pmatrix} 1/3 & 1/2 & 1/6 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/6 & 0 & 0 & 0 \\ 1/3 & 1/2 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 \end{pmatrix}$$

and the update of the second component according to

$$P_2 = \begin{pmatrix} 4/7 & 0 & 0 & 3/7 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 2/5 & 0 & 0 & 3/5 \\ 4/7 & 0 & 0 & 3/7 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 2/5 & 0 & 0 & 3/5 \end{pmatrix}$$