Ejercicios para el control 1
CC50Q - Teoría de la Información y Redes Neuronales

Prof: Pedro Ortega <peortega@dcc.uchile.cl>
Aux: Francisco Claude <fclaude@dcc.uchile.cl>

18 de agosto de 2005

Ejercicio 1

Si ud. dispone de la tabla de probabilidades conjuntas $P(x, y)$ de dos arreglos de probabilidad discretos X e Y, ¿en qué debe fijarse para determinar que existe dependencia? ¿Cómo cambia su criterio si a cambio posee la tabla de probabilidades condicionales $P(y|x)$?

Ejercicio 2

Una urna contiene K bolas, de las cuales B son negras y $W = K - B$ son blancas. Fred escoge una bola al azar de la urna y la reemplaza, N veces.

1. ¿Cual es la *distribución de probabilidad* de del número de veces de escoger bolas negras n_B?
2. ¿Cuál es la esperanza de n_B? ¿Cuál es la varianza de n_B? ¿Cuál es la desviación estándar de n_B? Entregue una respuesta numérica para el caso $N = 5$ con $B = 2$ y $K = 10$.

Ejercicio 3

Una urna contiene K bolas, B negras y $W = K - B$ blancas. Definimos la fracción $f_B \equiv B/K$. Fred escoge N de la urna igual que en el ejercicio anterior, obteniendo n_B negras, y calcula la cantidad

$$z = \frac{(n_B - f_B N)^2}{N f_B (1 - f_B)}$$

¿Cuál es la esperanza de z? Para el caso $N = 5$ y $f_B = 1/5$, cuál es la distribución de probabilidad de z? Cuál es la probabilidad de que $z < 1$?

Ejercicio 4

Sea $p_a = 0{,}1$, $p_b = 0{,}2$ y $p_c = 0{,}7$. Sea $f(a) = 10$, $f(b) = 5$, $f(c) = 10/7$. ¿Qué es $\mathcal{E}[f(x)]$? ¿Qué es $\mathcal{E}[1/P(x)]$?

Ejercicio 5

Para un arreglo arbitrario, ¿qué es $\mathcal{E}[1/P(x)]$?

Ejercicio 6

Muestre que si x es una variable binaria, se cumple:

$$\ln\left(\frac{P(x=1|y)}{P(x=0|y)}\right) = \ln\left(\frac{P(y|x=1)}{P(y|x=0)}\right) + \ln\left(\frac{P(x=1)}{P(x=0)}\right)$$

Ejercicio 7

Hay tres puertas, 1,2 y 3. Un premio ha sido escondido tras una de las puertas. Ud. puede escoger una puerta. Inicialmente la puerta que Ud. escogió no es abierta. A cambio, el anfitrión del juego abrirá una de las dos restantes, *de manera de no revelar la puerta con el premio.* Tras ésto, Ud. puede cambiar su elección o mantenerla igual.

Imagínese que un participante escoge primero a la puerta 1. Luego, el anfitrión abre la puerta 3, mostrando que ahí no estaba el premio. ¿El participante debería (a) mantener su elección, (b) cambiarse a la puerta 2 o (c) no importa?

Ejercicio 8

Para el mismo juego anterior, imagínese que justo cuando el anfitrión estaba a punto de abrir una de las puertas, un terremoto mueve todo el edificio, y una de las puertas se abre aleatoriamente. Resultó se la puerta 3, tras la cual no estaba el premio. El participante había escogido la puerta 1.

El anfitrión dice: 'bueno, hagamos como si yo mismo hubiese abierto esa puerta y sigamos adelante'. Nuevamente, ¿el participante debería (a) mantener su elección, (b) cambiarse a la puerta 2 o (c) no importa? Asuma que el premio fue ubicado al azar, que el anfitrión no sabe donde está, y que la puerta 3 se abrió al azar.

Ejercicio 9

Los habitantes de ua isla mienten $\frac{2}{3}$ de las veces y el tercio restante dicen la verdad. Si luego de escuchar una frase de uno de los habitantes, preguntamos si lo anterior era cierto y nos responde que sí. ¿Cuáll es la probabilidad de que la frase haya sido cierta?

Ejercicio 10

Una bolsa contiene un chip que es o blanco o negro. Se agrega un chip blanco a la bolsa y se mezcla. ¿Cuál es la probabilidad de escoger un chip blanco?

Ejercicio 11

Ud. se muda a una nueva casa. El teléfono está conectado, y Ud. está casi seguro de que el número es 277'34'46, pero no completamente. Ud. realiza un experimento: levanta el auricular y marca 277'34'46, obteniendo la señal de ocupado. ¿Está Ud. ahora más seguro de su número de teléfono? ¿Si está más seguro, cuánto más?

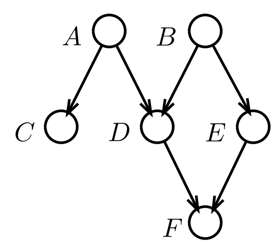
Ejercicio 12

Se tienen 2 dados equilibrados que pueden tomar los valores $\{1, 2, 3, 4, 5, 6\}$ cada uno.

- Describa el arreglo de probabilidad de la variable x, donde x representa la suma de los dos dados luego de lanzarlos.
- Asigne nuevos valores a los dados, usando el conjunto $\{0, 1, 2, 3, 4, 5, 6\}$, de forma que la suma de ambos tenga una distribución uniforme entre 1 y 12.

Ejercicio 13

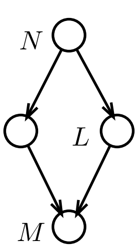
a) Suponga que Ud. tiene un problema que involucra seis variables binarias para su descripción, i.e. la probabilidad conjunta $P(A, B, C, D, E, F)$ describe al problema. ¿Cuántas probabilidades $p_i = P(A, B, C, D, E, F)$ necesita saber? b) Suponga ahora que Ud. modela al problema por medio de la red Bayesiana ilustrada a continuación. Encuentre una expresión simplificada de la probabilidad conjunta $P(A, B, C, D, E, F)$. ¿Cuántas probabilidades necesita ahora para describir al mismo problema?



c) Si Ud. conoce el valor de las variables C y E, encuentre una expresión para calcular la probabilidad de la variable A.

Ejercicio 14

a) Considere la red Bayesiana ilustrada a continuación, donde N='Está nublado', C='Está chispeando', L='Llueve', M='El pasto está mojado'.



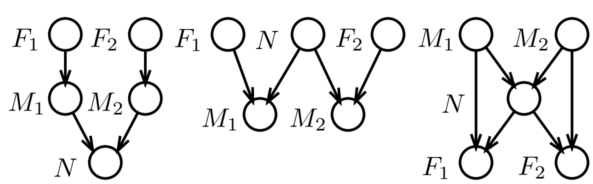
Ud. dispone de las siguientes tablas de probabilidades:

$P(N) = 0{,}50$	
$P(C\|N=1) = 0{,}10,$	$P(C\|N=0) = 0{,}50$
$P(L\|N=1) = 0{,}80,$	$P(L\|N=0) = 0{,}20$
$P(M\|C=1, L=1) = 0{,}99,$	$P(M\|C=1, L=0) = 0{,}90$
$P(M\|C=0, L=1) = 0{,}90,$	$P(M\|C=0, L=0) = 0{,}00$

Si el pasto está mojado, ¿cuál es la probabilidad de que esté nublado? Si luego observa que está chispeando, ¿cómo cambia la probabilidad de que esté nublado?

Ejercicio 15

Dos astrónomos, en distintos lugares del mundo, realizan medidas M_1 y M_2 del número de estrellas N de una región muy reducida del cielo, utilizando sus telescopios. Normalmente, existe un probabilidad pequeña de cometer hasta un error (1 estrella) en la medición. Además, existe una probabilidad menor de que los telescopios estén desenfocados (F_1 y F_2 respectivamente), en cuyos casos el científico medirá por lo menos tres estrellas menos de las que hay.



a) ¿Cual de las tres redes Bayesianas representa correctamente la situación (no es necesariamente la más eficiente)? b) Construya una tabla de probabilidades condicionales razonables para $P(M_1|N)$. (Para este problema, considere que sólo los valores 1, 2 y 3 son posibles). c) Suponga que $M_1 = 1$ y $M_3 = 3$. ¿Cuáles son los posibles números de estrellas N? d) ¿Cuál de las anteriores es la más probable?

Ejercicio 16

Una moneda es lanzada 250 veces, de los cuales 140 fueron cara y 110 sello. Compare las hipotesis moneda sesgada v/s moneda insesgada.

Ejercicio 17

Supongama que la distribución a priori para la proporción p de bebidas de una máquina despachadora que las derrama es:

p	0.05	0.10	0.15
$f(p)$	0.3	0.5	0.2

- Calcule la distribución a posteriori para p.
- Estime p, usando el criterio de máxima verosimilitud. ¿Qué tanto influye el conocimiento a priori que teníamos de p?

Ejercicio 18

Desarrolle el mismo problema anterior, asumiendo la distribución a priori:

$$f(p = 10), \quad 0{,}05 < p < 0{,}15$$

Ejercicio 19

N datos $\{x_n\}$ son obtenidos de N distribuciones de probabilidad. Todas las distribuciones son gaussianas con media μ común pero con varianzas σ_n desconocidas. ¿Cuáles son los parámetros de máxima verosimilitud $\mu, \{\sigma_n\}$, datos los datos? Por ejemplo, siete científicos (A,B,C,D,E,F,G) con habilidades de medición muy disímiles miden μ. Algunos de ellos harán un trabajo más preciso que otros (con σ_n bajo), y algunos de ellos tienen muy baja precisión. La tabla siguiente muestra las estimaciones x_n de μ obtenidas :

Científico	x_n
A	-27.020
B	3.570
C	8.191
D	9.898
E	9.603
F	9.945
G	10.056

Intuitivamente, ¿cuáles de estos científicos son malos medidores? Compare su intuición contra el resultado obtenido por máxima verosimilitud.

Ejercicio 20

Un set de variables aleatorias x provienen en forma independiente de una misma distribución de probabilidad $P(x)$. Según el modelo $\mathcal{H}_0$, $P(x)$ es una distribución de probabilidad uniforme:

$$P(x|\mathcal{H}_0) = \frac{1}{2} \qquad x \in (-1, 1)$$

Según el modelo $\mathcal{H}_1$, $P(x)$ es una distribución no-uniforme con un parámetro desconocido $m \in (-1, 1)$:

$$P(x|m, \mathcal{H}_1) = \frac{1}{2}(1 + mx) \quad x \in (-1, 1)$$

Dados los datos $D = \{0.3, 0.5, 0.7, 0.8, 0.9\}$, ¿cuál es la evidencia para los modelos $\mathcal{H}_0$ y $\mathcal{H}_1$ (evidencia $= P(\mathcal{H}|D)$)?

Ejercicio 21

Se piensa que los datos (x, t) provienen de una línea recta. El experimentador escoge x, y t posee una distribución gaussiana centrada en

$$y = w_0 + w_1 x$$

con varianza σ_ν^2 (con distorsión vertical). Según el modelo $\mathcal{H}_1$, la línea es horizontal, i.e. $w_1 = 0$. Según el modelo $\mathcal{H}_2$, w_1 es un parámetro con una distribución a priori normal $\mathcal{N}(0, 1)$. Ambos modelos le asignan la misma distribución a priori $\mathcal{N}(0, 1)$ a w_0. Dados los datos $D\{(-8, 8), (-2, 10), (6, 11)\}$, y suponiendo que el nivel de ruido es $\sigma_\nu = 1$, ¿cuál es la evidencia para cada modelo?

Ejercicio 22

Ud. posee una serie de observaciones $D = \{(x_i, y_i)\}_{i=1}^n$ que desea *aproximar* con un polinomio de grado m:

$$y \backsim \mathcal{N}\left(\sum_{k=0}^{m} w_k x^k, \sigma\right)$$

donde los parámetros del modelo son los coeficientes del polinomio $\{w_k\}_{k=0}^m$ y la precisión σ. Usando el método de máxima verosimilitud, ¿cuánto valen la función verosimilitud y σ para los casos $m > n$ y $m < n$? Encuentre valores exactos cuando sea posible. Explique si existe un modelo mejor que el otro.

Ejercicio 23

Ud. posee un conjunto de datos $D = \{x_i\}_{i=1}^n$ muestreados de una misma variable aleatoria y un *modelo generador* de estos datos $f(x|\vec{w})$. Ud. quiere modelar la densidad de probabilidad real de los datos por medio de su función f, por lo que la ha diseñado de manera que

$$\int f(x|\vec{w})dx = 1$$

para cualquier elección del vector de parámetros $\vec{w}$. Este último tiene 15 componentes, i.e. $\vec{w} \in \mathbb{R}^{15}$. a) Si su objetivo es hallar una buena elección de $\vec{w}$, escriba la solución Bayesiana a este problema. b) Ahora, suponga que la solución es tan compleja que no puede calcular una solución analítica, por lo que ha optado utilizar métodos computacionales para calcular la integral (método de Montecarlo). Para esto, Ud. muestrea repetitivamente un vector $\vec{w}$ de su distribución de probabilidad a priori $P(\vec{w})$, evalúa, y utiliza el resultado para aproximar el resultado Bayesiano (el histograma resultante). Explique si este método tiene sentido, y cuáles son los peligros de su utilización. c) ¿Ud. lo haría? Estime cuántas iteraciones del método de Montecarlo serían necesarias.

Ejercicio 24

Dos dados insesgados se lanzan repetitivamente: el resultado de la tirada es la suma de ambos dados. Juan, quien dice que el 6 y el 8 son sus números de la suerte, incluso apuesta dinero de que se lanzará el 6 antes del 7. ¿Si Ud. fuese un jugador, apostaría? ¿Cuál es su probabilidad de ganar/perder? Si Juan ahora decide apostar a que 8 saldrá antes del 7, ¿apostaría Ud.?

Juan se vuelve más osado, y decide combinar ambas apuestas en una sola: si logra salir un 6 y un 8 antes de que salgan dos 7, gana. ¿Apostaría Ud.? ¿Cuál es su probabilidad de ganar?