Estadística Actuarial (751)
Cátedra: Landro
Curso: Landro – Del Rosso
1er Cuatrimestre de 2018

Práctica 1 de R

Contenido

Ejercicio 1

Ejercicio de aplicación desarrollado por las Profesoras Patricia GIRIMONTE y Claudia MOLINARI

Los siguientes datos corresponden a un ejemplo hipotético de 10 empleados del área administrativa de una empresa.

Las variables observadas para cada empleado son: Edad (años), Sexo, Sueldo y Antigüedad (en años al 31 de octubre)

Ingresar los siguientes datos como una matriz de nombre “Datos.Empleados” **donde cada fila identifique a cada individuo**

N° Empleado	1	2	3	4	5	6	7	8	9	10
Edad	51	44	55	47	56	54	53	45	48	55
Sexo	M	F	F	F	M	F	M	F	M	M
Sueldo	7200	6700	5000	5200	7500	6300	8700	6100	9200	9000
Antigüedad	5	6	4	5	7	4	8	6	10	9

a. Definir un vector con los datos de cada uno de los empleados.
b. Con los vectores del punto a) arme una matriz de nombre “Datos.Empleados” donde cada fila identifique a cada empleado.
c. Escriba los comandos adecuado para,
    1. Visualizar en pantalla la cantidad de individuos y variables ingresadas.
    2. Ponerles nombre a las variables
    3. Visualizar en pantalla el Sueldo del individuo n°5 ingresado
    4. Visualizar en pantalla los datos completos del individuo n° 8
    5. Visualizar en pantalla los Sueldos de los hombres.
    6. Visualizar en pantalla la Antigüedad de las mujeres.
    7. Visualizar en pantalla el Sueldo de los empleados de más de 50 años.
    8. Generar un vector con el valor de la variable (Sueldo/Antigüedad) de cada individuo.
    9. Adicionar el vector generado en h) como una nueva columna de la matriz “Datos. Empleados”
d. Transformar la matriz “Datos.Empleados” en un data frame.
e. Guardar el data frame como un archivo csv.

Ejercicio 2

Las edades de un grupo de amigos son 27, 23, 29, 24 y 31 años. Crear un vector “edades” con estos datos y calcular su media, de forma que la salida se guarde en un fichero llamado “amigos”. Volver a calcularla, pero de forma que ahora el resultado se exhiba en la pantalla.

¿Cómo se obtiene el archivo de texto con los últimos comandos ejecutados?

Ejercicio 3

En clase vimos diferentes formas de definir vectores con R. Supongamos que queremos definir el vector $x = (1,2,3,4,5)$. Comprobar que las siguientes formas son equivalentes,

$> x < -c(1,2,3,4,5)$

$> x < -1{:}\,5$

$> x - seq(1,5)$

Ejercicio 4

Definir el vector $y = (1,3,5,7)$ utilizando la función $c()$.

¿Cómo lo harían con la función $seq()$?.

Recuerden que si tienen alguna duda sobre cómo se definen las funciones siempre pueden consultar la ayuda.

Ejercicio 5

Definir los siguientes vectores.

Intenta hacerlo de diferentes formas.

$x = (8,7,6,5)$

$y = (3,3,3,3,3,3,3,3,2,2)$

$z = (1,1.75,2.5,3.25,4)$

Ejercicio 6

Aunque pensamos en vectores como conjuntos de números, un vector en R no es más que celdas contiguas conteniendo datos. Estos datos deben ser del mismo tipo, pero no necesariamente números.

Podemos construir así vectores de tipo *logical* o vectores de tipo *character* entre otros. Por ejemplo, hemos creado el vector chica.

El resultado es,

$> profesores$

[1] "Alberto" "Daniel" "Pablo" "Eduardo"

¿Cómo ha sido definido dicho vector?

Ejercicio 7

En muchas ocasiones nos interesa hacer referencia a determinadas componentes de un vector. En clase vimos que para ello utilizamos los corchetes [ ] .

Crear el vector $x = (2, -5, 4, 6, -2, 8)$. A partir de dicho vector definir,

a) $y = (2, 4, 6, 8)$. Así definido y es el vector formado por las componentes positivas de x.

b) $z = (-5, -2)$. Así definido z es el vector formado por las componentes negativas de x.

c) $v = (-5, 4, 6, -2, 8)$. Así definido v es el vector x eliminada la primera componente.

d) $w = (2, 4, -2)$. Así definido w es el vector x tomando las componentes impares

Ejercicio 8

Sabemos que para sumar vectores éstos deben tener la misma longitud. Sin embargo R trabaja de manera distinta.

Definir los vectores $x = (1,2,3,4,5,6), y = (7,8), z = (9,10,11,12)$.

Calcular,

a) $x + x$

b) $x + y$. ¿Qué hace R?

c) $x + z$. Ahora R da un *warning* pero aun así nos da un resultado. ¿Cómo lo ha calculado?

Ejercicio 9

Define el vector $x = (1,2,3,4,5,6)$.

A partir de dicho vector se han construido las matrices $m1, m2, m3$ y $m4$.

$> m1$

	[,1]	[,2]	[,3]
[1,]	1	3	5
[2,]	2	4	6

$> m2$

	[,1]	[,2]
[1,]	1	4
[2,]	2	5
[3,]	3	6

$> m3$

	[, 1]	[, 2]	[, 3]
[1,]	1	2	3
[2,]	4	5	6

$> m4$

	[, 1]	[, 2]	[, 3]
[1,]	1	4	1
[2,]	2	5	2
[3,]	3	6	3

Todas las matrices se han definido a partir de $matrix(x, \dots)$.

Intentar reproducir el código necesario para obtener cada una de ellas. Recordar que pueden consultar la ayuda

Ejercicio 10

¿Qué ocurre cuando definimos una matriz en R y sólo especificamos el número de filas o el número de columnas?

¿Qué ocurre cuando los datos no se corresponden con la dimensión de la matriz que queremos definir?

Comprobarlo ejecutando los siguientes comandos,

$> matrix(1{:}6, nrow = 2)$

$> matrix(1{:}6, nrow = 4)$

$> matrix(1{:}6, nrow = 4, ncol = 4)$

Ejercicio 11

¿Cuál es la diferencia entre $*$, $\% * \%$ y $outer()$? Comprobarlo con las matrices

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} \quad y \quad B = \begin{pmatrix} 3 \\ 8 \end{pmatrix}$$

Ejercicio 12

Sean A una matriz 2 × 3, B una matriz 3 × 4 y C una matriz 2 × 3.

¿De qué tipo y dimensión serán los objetos obtenidos de los siguientes comandos de R?

¿Alguno de los comandos produce mensajes de error? ¿Por qué?

a) $A * B$

b) $outer(A, B)$

c) $A + 2$

d) $A\% * \%B$

e) $\exp(B)$. Nota: $\exp()$ es la función exponencial.

f) $A * C$

g) $A\% * \%C$

Ejercicio 13

Crear en R el vector $x = (1,2,3,4,5,6)$ y definir a partir de él las siguientes matrices,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad D = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$$

Ejercicio 14 (Cadena de Markov)

***Ejercicio extraído de la Sección 1 del libro "Matemática para Economistas* con Microsoft Excel y MATLAB" de los autores Alicia BERNARDELLO, María José BIANCO, María Teresa CASPARRI, Javier García FRONTI y Susana OLIVERA DE MARZANA.**

Supongamos un mercado duopolista, en el que las empresas A y B fabrican la totalidad de cierto producto. Mediante estudios de mercado se ha arribado a las siguientes conclusiones,

1) El 50% de los consumidores que compran durante un mes el producto fabricado por la empresa A volverá a hacerlo en el mes siguiente, pero el resto comprará el fabricado por la empresa B.
2) El 25% de los consumidores que en un mes compran el producto fabricado por la empresa B volverá a hacerlo así el mes siguiente, pero el resto cambiará el fabricado por la empresa A.
3) En este momento la empresa A tiene el 40% del mercado y la empresa B el 60% restante.

El vector de probabilidad que representa la situación inicial se denomina "vector de estado inicial"; en este caso es $p_0 = (0{,}4; 0{,}60)$.

La matriz de transición es $P = \begin{pmatrix} 0{,}50 & 0{,}50 \\ 0{,}75 & 0{,}25 \end{pmatrix}$

Se solicita efectuar los siguientes cálculos en R.

a. Crear un vector que represente el estado inicial.
b. Crear una matriz que represente las transiciones de un estado a otro.
c. Calcular los vectores que representen las situaciones en los momentos t_1 (al finalizar el primer mes), t_2, t_5 y t_n (para éste último generalícelo en R).
d. Verificar con el comando $sum()$ que dichos vectores cumplen con ser vectores de probabilidad o estocástico y con el comando $apply()$ que la matriz de transición es estocástica.
e. Calcular el vector t que representa el estado firme o fijo de la Cadena Markov.

Ejercicio 15 (Cadena de Markov)

Ejercicio extraído de la Sección 1 del libro "Matemática para Economistas **con Microsoft Excel y MATLAB" de los autores Alicia BERNARDELLO, María José BIANCO, María Teresa CASPARRI, Javier García FRONTI y Susana OLIVERA DE MARZANA.**

Supongamos la siguiente matriz de transición que posee dos estados absorbentes $S_3\ y\ S_4$

$$P = \begin{pmatrix} 0{,}4 & 0{,}3 & 0{,}2 & 0{,}1 \\ 0{,}1 & 0{,}1 & 0{,}6 & 0{,}2 \\ 0{,}0 & 0{,}0 & 1{,}0 & 0{,}0 \\ 0{,}0 & 0{,}0 & 0{,}0 & 1{,}0 \end{pmatrix}$$

Si el estado inicial es S_3 los vectores de estados sucesivos serán siempre los mismos,

$$p_0 = (0{,}0{,}1{,}0), p_1 = (0{,}0{,}1{,}0) \ldots\ p_n = (0{,}0{,}1{,}0)$$

Si el estado inicial es S_4 los vectores de estados sucesivos serán siempre los mismos,

$$p_0 = (0{,}0{,}0{,}1), p_1 = (0{,}0{,}0{,}1) \ldots\ p_n = (0{,}0{,}0{,}1)$$

Pero si el estado inicial es S_1 los estados sucesivos serán,

$$p_0 = (1{,}0{,}0{,}0), p_1 = (0{,}4; 0{,}3; 0{,}2{,}0{,}1), p_{20} = (0{,}0000003; 0{,}0000003; 0{,}7058819; 0{,}2941175)$$

Si el estado inicial fuera S_2 los estados sucesivos serán,

$$p_0 = (0{,}1{,}0{,}0), p_1 = (0{,}1; 0{,}1; 0{,}6{,}0{,}2), p_{20} = (0{,}00000009; 0{,}00000007; 0{,}745098; 0{,}254902)$$

Para calcular directamente la probabilidad a largo plazo de que partiendo de un estado no absorbente se arribé a uno u otro de los estados absorbentes se construyen dos submatrices con las filas que corresponden a los estados no absorbentes: una con las columnas que corresponden a los estados absorbentes, que se denomina G y otra con las que corresponden a los no absorbentes denominada H. Es decir,

$$G = \begin{pmatrix} 0{,}2 & 0{,}1 \\ 0{,}6 & 0{,}2 \end{pmatrix} \qquad H = \begin{pmatrix} 0{,}4 & 0{,}3 \\ 0{,}1 & 0{,}1 \end{pmatrix}$$

Se solicita efectuar los siguientes cálculos en R con los comandos vistos en clase.

a. Calcular los anteriores vectores de estados considerando que se parte únicamente de un estado.
b. Verificar con el comando $sum()$ que dichos vectores cumplen con ser vectores de probabilidad o estocásticos y con el comando $apply()$ que la matriz de transición es estocástica.
c. Calcular la matriz $Q = (I - H)^{-1}$.
d. Calcular la matriz $R = Q * G$

Ejercicio 16

Crear una función que simplemente devuelva el área de un rectángulo tomando como argumentos, su base y su altura.