PROBABILIDADES Y ESTADÍSTICA (C)

PARCIAL DE LABORATORIO - 27/06/2013 - TURNO 1 -

NOTA: *Realizar el examen en un script de R cuyo nombre será los dos apellidos de la pareja y el turno en el que rinden (por ej: perez_alvarez_t1). Colocar apellido y nombre de cada integrante del grupo en la primera línea del script, junto con el turno en el que rinde (1,2 ó 3). En el mismo hacer todos los comentarios que sean necesarios para justificar adecuadamente los ejercicios (recuerde usar # para todos los comentarios o líneas no ejecutables). Indique el número de ej y letra del ítem que está resolviendo (por ej: #2)b)). Verifique que el script se pueda correr o ejecutar adecuadamente. Entregar el script por mail, cuyo subject o título deberá ser el mismo nombre que le puso al archivo, a las siguientes direcciones* `aferrari@dm.uba.ar`, `andreslemm@gmail.com`. *Al retirarse entregar la hoja de enunciados con firma y aclaración de cada integrante. El examen consiste en resolver solamente dos ejercicios de los tres. Pueden tener lo que les parezca útil a su disposición, usar la ayuda online del R, pero no pueden abrir el correo electrónico ni otra red social durante la resolución del mismo.*

1. El precio de una acción se mueve día a día dentro de los valores $\{1, 2, 3, 4\}$ (expresado en alguna unidad monetaria). Sea X_n = precio de la acción el día n. Supondremos que $\{X_n\}_{n\geq 1}$ es una cadena de Markov verificando: si $X_n = j$ con $j = 2$ o 3, luego $X_{n+1} = j - 1$ con probabilidad 0,2 y $X_{n+1} = j + 1$ con probabilidad 0,8. Si $X_n = 1$ entonces $X_{n+1} = 2$ con probabilidad 1. Finalmente, si $X_n = 4$ entonces $X_{n+1} = 3$ con probabilidad 1.

   a) Hacer un programa que tenga como input la matriz de transición de la cadena, una probabilidad inicial llamada p, un número natural n y devuelva $x_1, x_2 ......x_n$ realizaciones de la cadena de Markov arrancando con la probabilidad inicial p. Un ejemplo de p podría ser $p = (\frac{1}{2}, \frac{1}{2}, 0, 0)$. Fijar la semilla con el set. seed.

   b) Usando el item anterior aproxime a la probabilidad invariante.

2. Se dispone de dos canastos con pelotas de dos colores, cuyo contenido es el siguiente:

   canasto A: 4 pelotas rojas y 5 blancas.

   canasto B: 1 pelota roja y 3 blancas.

   Por otro lado un bolillero tiene 3 bolitas, dos de color verde y una azul. Un juego consiste en elegir una bola del bolillero, si sale verde se extrae una pelota del canasto A y se la coloca en B para luego sacar una pelota del canasto B y anotar el color de la misma. En caso contrario el proceso se hace a la inversa.

   a) Simular una realización del juego. Fijar la semilla con el set.seed. Construir una función que tenga como input un número natural n, el número de semilla y que devuelva n repeticiones independientes del juego.

   b) Usando el item anterior y el concepto de frecuencia relativa calcular en forma aproximada la probabilidad de obtener roja en el juego.

   c) Verifique los resultados obtenidos en el item anterior calculando en forma exacta la probabilidad de sacar una bola roja en el juego.

3. Sea X una variable aleatoria con densidad:

$$f_X(x) = \begin{cases} \frac{4}{3}x^{1/3} & 0 \leq x \leq 1 \\ 0 & \text{en otro caso.} \end{cases}$$

a) Realizar un gráfico de f usando los puntos $(x_i, f(x_i))_{i\in\{1,..,100\}}$, con x_i una partición uniforme del intevalo $[0, 1]$.

b) Sea F la distribución acumulada de X. Se pide generar realizaciones de $X_1, .., X_n$ independientes con distribución F. Es decir: armar un programa que tenga como input un número natural n y que genere n números aleatorios (en realidad pseudo aleatorios) para la distribución F. Fijar la semilla con el set.seed.

c) Usando el item anterior, para un n suficientemente grande, aproxime el valor $P(1/4 < X < 1/2)$ y comprarlo con el valor exacto.

d) Utilizar el programa construido en el item b) para generar muestras de tamaño $n = 100, 500, 1000$ y 5000. Para cada una de ellas dibujar un histograma de frecuencias razonable (es decir, de ser necesario modificar los parámetros del mismo para que resulte representativo). Observar cómo evoluciona el histograma a medida que el tamaño de la muestra aumenta. Sacar conclusiones.

PROBABILIDADES Y ESTADÍSTICA (C)

PARCIAL DE LABORATORIO - 27/06/2013 - TURNO 2 -

NOTA: *Realizar el examen en un script de R cuyo nombre será los dos apellidos de la pareja y el turno en el que rinden (por ej: perez_alvarez_t2). Colocar apellido y nombre de cada integrante del grupo en la primera línea del script, junto con el turno en el que rinde (1,2 ó 3). En el mismo hacer todos los comentarios que sean necesarios para justificar adecuadamente los ejercicios (recuerde usar # para todos los comentarios o líneas no ejecutables). Indique el número de ej y letra del ítem que está resolviendo (por ej: #2)b)). Verifique que el script se pueda correr o ejecutar adecuadamente. Entregar el script por mail, cuyo subject o título deberá ser el mismo nombre que le puso al archivo, a las siguientes direcciones* `molina_ julieta@yahoo. com. ar`, `mfragala@ dm. uba. ar`. *Al retirarse entregar la hoja de enunciados con firma y aclaración de cada integrante. El examen consiste en resolver solamente dos ejercicios de los tres. Pueden tener lo que les parezca útil a su disposición, usar la ayuda online del R, pero no pueden abrir el correo electrónico ni otra red social durante la resolución del mismo.*

1. El ascensor de un edificio con planta baja y dos pisos realiza viajes de uno a otro piso. Definimos

   X_n = piso en el que el ascensor finaliza el viaje n-ésimo.

   Supondremos que $\{X_n\}_{n\geq 1}$ sigue una cadena de Markov. Se sabe que la mitad de los viajes que parten del la planta baja se dirigen a cada uno de los otros dos pisos, mientras que si un viaje comienza en el primer piso, sólo el 25 % de las veces finaliza en el segundo. Por último, si un trayecto comienza en el segundo piso, siempre finaliza en la planta baja.

   a) Hacer un programa que tenga como input la matriz de transición de la cadena, una probabilidad inicial llamada p, un número natural n y devuelva $x_1, x_2......x_n$ realizaciones de la cadena de Markov arrancando con la probabilidad inicial p. Un ejemplo de p podría ser $p = (\frac{1}{2}, \frac{1}{2}, 0)$. Fijar la semilla con el set. seed.

   b) Usando el item anterior aproxime a la probabilidad invariante usando un valor suficientemente grande.

2. Un juego consiste en arrojar dos dados equilibrados y observar la suma obtenida. El objetivo del ejercicio es simular las probabilidades de que la suma sea 9 dado que:

   i) la suma es impar.

   ii) la suma es múltiplo de 3.

   Se pide:

   a) Simular una realización del juego fijando una semilla con el set.seed. Construir una función que tenga como input un número natural n y que devuelva n repeticiones independientes del juego.

   b) Mediante el item anterior y el concepto de frecuencia relativa, calcular en forma aproximada la probabilidad de que la suma sea 9 dado que la suma es impar. Fije una semilla usando set.seed.

   c) Mediante el item anterior y el concepto de frecuencia relativa, calcular en forma aproximada la probabilidad de que la suma sea 9 dado que la suma es múltiplo de tres. Fije una semilla usando set.seed.

d) Si tiene tiempo verifique los resultados obtenidos.

3. Sea X una variable aleatoria con densidad:

$$f(x) = \frac{1}{2\sqrt{3}}\sqrt{x} \;\; I_{[0,3]}(x).$$

a) Realizar un gráfico de f usando los puntos $(x_i, f(x_i))_{i\in\{1,..,100\}}$, con x_i una partición uniforme del intevalo $[0, 3]$.

b) Sea F la distribución acumulada de X. Se pide generar realizaciones de $X_1, .., X_n$ independientes con distribución F. Es decir: armar un programa que tenga como input un número natural n y que genere n números aleatorios (en realidad pseudo aleatorios) para la distribución F. Fijar la semilla con el set.seed.

c) Usando el item anterior, para un n suficientemente grande, aproxime el valor $P(1 < X < 2)$ y comprarlo con el valor exacto.

d) Utilizar el programa construido en el item b) para generar muestras de tamaño $n = 100, 500, 1000$ y 5000. Para cada una de ellas dibujar un histograma de frecuencias razonable (es decir, de ser necesario modificar los parámetros del mismo para que resulte representativo). Observar cómo evoluciona el histograma a medida que el tamaño de la muestra aumenta. Sacar conclusiones.

PROBABILIDADES Y ESTADÍSTICA (C)

Parcial de Laboratorio - 27/06/2013 - Turno 3 -

NOTA: *Realizar el examen en un script de R cuyo nombre será los dos apellidos de la pareja y el turno en el que rinden (por ej: perez_alvarez_t3). Colocar apellido y nombre de cada integrante del grupo en la primera línea del script, junto con el turno en el que rinde (1,2 ó 3). En el mismo hacer todos los comentarios que sean necesarios para justificar adecuadamente los ejercicios (recuerde usar # para todos los comentarios o líneas no ejecutables). Indique el número de ej y letra del ítem que está resolviendo (por ej: #2)b)). Verifique que el script se pueda correr o ejecutar adecuadamente. Entregar el script por mail, cuyo subject o título deberá ser el mismo nombre que le puso al archivo, a las siguientes direcciones* `molina_julieta@yahoo.com.ar`, `mfragala@dm.uba.ar`. *Al retirarse entregar la hoja de enunciados con firma y aclaración de cada integrante. El examen consiste en resolver solamente dos ejercicios de los tres. Pueden tener lo que les parezca útil a su disposición, usar la ayuda online del R, pero no pueden abrir el correo electrónico ni otra red social durante la resolución del mismo. El único ítem que deben entregar en la hoja de enunciados es el 3)a). Intenten realizar códigos prolijos y cuando sea necesario justificar claramente.*

1. En Colombia existen 3 operadores principales de telefonía móvil como lo son Tigo, Comcel y Movistar. Los proporciones actuales que tiene cada operador en el mercado son para Tigo 0.4, para Comcel 0.25 y para Movistar 0.35.

   Se cuenta con la siguiente información: un usuario actualmente de Tigo tiene una probabilidad de permanecer en Tigo el próximo año de 0.60, de pasar a Comcel 0.2 y de pasarse a Movistar de 0.2. Si en la actualidad el usuario es cliente de Comcel tiene una probabilidad de mantenerse en Comcel el próximo año de 0.5, que se cambie a Tigo 0.3 y que se pase a Movistar de 0.2. Si el usuario es cliente en la actualidad de Movistar la probabilidad que permanezca en movistar el próximo año es de 0.4, que se cambie a Tigo de 0.3 y a Comcel de 0.3.

   a) Actualmente Juan es cliente de Comcel. Calcular la probabilidad de que dentro de 1 año, 3 años y 20 años siga siendo cliente de la misma empresa.

   b) Calcular los proporciones de clientes dentro de 1 año, 3 años y 20 años que tendrá cada operador en el mercado.

   c) Construir una función que tenga como input la matriz de transición del sistema, la distribución inicial, un número natural n y devuelva las proporciones de clientes que tendrá cada operador en n años. Para los datos de nuestro problema, ¿qué ocurre con estas proporciones para n suficientemente grande ($n = 20, n = 50, n = 100$)?

   d) Simule las primeras 10000 realizaciones de la Cadena de Markov comenzando en el estado Movistar y utilizarlo para calcular aproximadamente la medida invariante. Fijar la semilla con el set.seed. Comparar este resultado con el obtenido en el ítem anterior.

   e) ¿Encuentra diferencias significativas si en vez de comenzar en Movistar comienza en Tigo o en Comcel? A partir de lo que observó computacionalmente, ¿la medida invariante estaría dependiendo del estado inicial?

2. Un juego consiste en arrojar dos dados equilibrados. El objetivo del ejercicio es simular las probabilidades de que la suma sea 7 dado que:

   i) el número del segundo dado es par.

ii) la suma es impar.

Se pide:

a) Mediante el concepto de frecuencia relativa, calcular en forma aproximada la probabilidad de que la suma sea 7 dado que la suma es impar. Para ellos construir una función que tenga como input las n repeticiones del juego y el valor de la semilla.

b) Mediante el concepto de frecuencia relativa, calcular en forma aproximada la probabilidad de que la suma sea 7 dado que el número del segundo dado es par. Para ellos construir una función que tenga como input las n repeticiones del juego y el valor de la semilla.

c) Verifique los resultados obtenidos calculando la probabilidad exacta. Entreguar solamente el resultado de la probabilidad (en fracción) y no su desarrollo.

3. Sea $U_1, \ldots, U_n$ una muestra aleatoria con distribución uniforme en $[0, 1]$ Sea f la densidad de una variable aleatoria X con distribución normal estándar.

   a) Se define $I_n = \frac{1}{n} \sum_{i=1}^{n} f(U_i)$. Explicar por qué I_n puede utilizarse para calcular en forma aproximada $\int_0^1 f(x)dx$.

   b) Utilizar el ítem anterior para calcular en forma aproximada $\int_0^1 f(x)dx$ a partir de una muestra aleatoria de tamaño $n = 1000$. Fijar la semilla con el set.seed.

   c) Extender el método utilizado en el ítem b) para aproximar $\int_a^b f(x)dx$, con $a < b$. Es decir construir una función que tenga de input los valores a, b, el tamaño de muestra n, el valor de la semilla y que devuelva el valor aproximado de dicha integral.

   d) Calcular aproximadamente $P(1 < X < 3)$ y compararlo con el valor exacto (o mejor dicho el valor que da el R como exacto).