Probabilidad y Estadística (93.24)
Trabajo Práctico N$^{\underline{o}}$ 6

Procesos Estocásticos: Procesos de Poisson y Cadenas de Markov

1. **La caminata aleatoria (o random walk) simétrica** Supongamos el siguiente juego: se lanza una moneda en forma reiterada y el jugador gana 1 peso si sale cara y pierde 1 peso si sale ceca.

   a) Se define el proceso estocástico $X_n : n \in \mathbf{N}$ donde X_n es el dinero que tiene el jugador al cabo de n jugadas (se supone $X_0 = 0$). Definir el recorrido de X_n y el espacio de estados del proceso (válido para todo valor de n).

   b) Obtener la distribución de probabilidades de X_n.

   c) Expresar la variable X_n como una suma de variables aleatorias independientes e idénticamente distribuidas y calcular $\mathrm{E}[X_n]$ y $\mathrm{Var}(X_n)$.

   Nota: Se llama caminata aleatoria porque, en vez de un juego, se puede pensar en la caminata de un individuo que se mueve un paso a la izquierda o a la derecha de acuerdo al resultado de lanzar una moneda.

   El siguiente código de *Octave* simula el proceso (se muestran M *realizaciones del proceso*y N instantes posteriores al inicial):

```
N = 100; M = 10;
for i=1:M
 Z(1,i) = 0;
 for k=1:N
 r = rand;s = -(r<1/2)+(r>=1/2);
 Z(k+1,i) = Z(k,i)+s;
 end
 end
 plot(Z,'o-')
```

2. **La caminata aleatoria (ó random walk) general** Extienda el ejercicio anterior, pero esta vez asumiendo que la moneda puede estar *cargada*, es decir, la probabilidad de que salga cara es $p \in (0, 1)$.

   a) Obtener la distribución de probabilidades de X_n.

   b) Calcular $\mathrm{E}[X_n]$ y $\mathrm{Var}(X_n)$.

3. Suponga un proceso estocástico en tiempo discreto con espacio de estados discreto definido de la siguiente manera:

$$X_{n+1} = X_n + Z_n$$

para n tomando los valores $0, 1, 2, \dots,$ (el *tiempo* discreto) y con $X_0 = 0$. Las variables aleatorias Z_n se suponen independientes, e igualmente distribuidas con recorrido $\{-2\,, -1\,, 0\,, 1\,, 2\}$ con función de probabilidad $p_Z(z) = P(Z = z)$ tal que $p_Z(-2) = p_Z(2) = 0.1$, $p_Z(-1) = p_Z(1) = 0.25$ y $p_Z(0) = 0.3$.

a) Obtener la distribución de probabilidades de X_n para n tomando los valores 1, 2 y 3.
b) Calcular $\mathrm{E}[X_n]$ y $\mathrm{Var}(X_n)$.

El siguiente código de *Octave* simula el proceso (se muestran M *realizaciones del proceso* y N instantes posteriores al inicial):

```
M=40; N=100;
X(1,1:M)=zeros(1,M);
S=[-2 -1 0 1 2];
P=[0.1 0.25 0.3 0.25.1];
Z=discrete_rnd (S,P,N,M);
for k=1:N
X(k+1,:)=X(k,:)+Z(k,:);
end
plot(X,'o-')
```

4. **La caminata aleatoria (o random walk) Gaussiana** Otra extensión de los ejercicios anteriores es cuando la cantidad perdida o ganada en cada jugada es una variable aleatoria continua. En particular, supongamos que la cantidad ganada en la jugada k-ésima es una variable aleatoria $G_k \sim N(0, 1)$. Más aún, asuma que las G_k son variables i.i.d. Se define el proceso estocástico $X_n : n \in \mathbf{N}$ donde X_n es el dinero que tiene el jugador al cabo de n jugadas (se supone $X_0 = 0$).

a) Definir el recorrido de X_n y el espacio de estados del proceso (válido para todo valor de n).
b) Obtener la distribución de probabilidades de X_n.
c) Calcular $\mathrm{E}[X_n]$ y $\mathrm{Var}(X_n)$.

Nota: En términos de una *caminata*, la longitud de los pasos dados varía en cada instante. Una forma intuitiva de pensarlo es como una partícula liviana suspendida en un líquido. Debido al movimiento de las moléculas del líquido, la partícula sufre continuos golpes o choques que cambian su trayectoria. Si discretizamos el tiempo (lo dividimos en una secuencia de instantes equiespaciados), podemos pensar que los golpes llevan a la partícula a pegar saltos de longitud variable. El resultado es algo parecido a lo que se conoce como *movimiento Browniano*, un proceso estocástico que tiene aplicaciones muy importantes en varios campos de la ciencia y de las finanzas.

El siguiente código de *Octave* simula el proceso (se muestran M *realizaciones del proceso* y N instantes posteriores al inicial):

```
X(1,1:40)=zeros(1,40);
for k=1:100
X(k+1,:)=X(k,:)+randn(1,40);
end
plot(X,'-o')
```

5. Se sabe que durante ciertas horas del día las llamadas telefónicas a una central están distribuidas al azar según un proceso de Poisson con un promedio de 4 llamadas por minuto. Calcular la probabilidad de que:

   a) transcurran dos minutos sin llamadas,

   b) en un minuto haya por lo menos dos llamadas,

   c) en tres minutos se produzcan exactamente 10 llamadas.

   d) en los próximos tres intervalos consecutivos de 3 minutos se produzcan 10 llamadas en cada uno.

   e) el tiempo entre dos llamadas consecutivas supere 30 segundos.

   f) el tiempo que transcurre hasta que ocurran 3 llamadas sea inferior a 1 minuto. Tenga en cuenta que para que el tiempo hasta que ocurran m eventos de un proceso de Poisson de parámetro λ sea inferior a t es equivalente a que el número de eventos en un intervalo de longitud t sea mayor o igual a m.

6. El número de partículas emitidas por una sustancia radiactiva en el tiempo obedece a un proceso de Poisson. Supongamos que la emisión ocurre a razón de 75 partículas por minuto.

   a) Determinar la probabilidad de que durante 20 segundos sean emitidas exactamente 20 partículas.

   b) Representar gráficamente la función de probabilidad de la variable aleatoria X: número de partículas emitidas en 20 segundos.

   c) Si en dos minutos se registraron 100 partículas, ¿cuál es la probabilidad de que en los próximos 45 segundos se registren 50 partículas más?

7. Se ha encontrado que el número de fallas de los subsistemas de un sistema dado puede considerarse un proceso de Poisson con un promedio de una falla de un subsistema cada 100 horas.

   a) Se inicia cierto proceso que requerirá que el sistema opere durante 200 horas. Calcular la probabilidad de que el proceso pueda ser completado con éxito si se supone que el sistema está operante si como máximo fallan 4 subsistemas.

   b) Suponga que en una situación especial se requiere operar uno de estos subsistemas. En cuanto falla se cambia por otro y asi cada vez que falla el que esta funcionando se reemplaza por otro. Se supone despreciable el tiempo que demora sacar el fallado y cambiarlo por el relevo. Se tienen 10 de estos subsistemas. ¿Cuál es la probabilidad de que ese stock pueda servir para cubrir un tiempo de servicio total de al menos 1500 hs.?

8. Suponga que en cierto banco se atiende, en promedio durante una parte del día, a cuatro clientes cada seis minutos según un proceso de Poisson. Calcular la probabilidad de que:

   a) puedan atenderse a seis o más clientes en seis minutos;
   b) se empleen más de tres minutos en atender a un cliente;
   c) el tiempo de atención a un cliente esté comprendido entre dos y cuatro minutos;
   d) el tiempo que insuma atender 10 clientes sea menor a 10 minutos.

9. El tiempo entre arribos de clientes, durante la mañana de un día normal, a una estación de servicio se puede considerar una variable aleatoria con distribución exponencial de media 5 minutos.

   a) Describir la distribución de probabilidades del número de clientes que llegan en una hora.
   b) Calcular la probabilidad de que en 30 minutos lleguen mas de 5 clientes.
   c) Calcular la probabilidad de que el intervalo de tiempo entre las llegadas de los clientes décimo y undécimo exceda 10 minutos.
   d) Calcular el valor esperado y la varianza del tiempo transcurrido hasta que llega el décimo cliente desde que comenzó el servicio.
   (Nota: Si se supone que el tiempo transcurrido es una suma de variables aleatorias independientes entonces se cumple que la varianza de esa suma es la suma de las varianzas de los sumandos).
   e) Usando la aproximación normal de la distribución de Poisson calcular la probabilidad de que el número de clientes atendidos en 6 horas exceda 90.

10. Los impulsos de ruido en una línea telefónica siguen un proceso de Poisson de parámetro λ (en impulsos por segundo).

    a) Determinar la probabilidad de que no haya ruido durante t segundos.
    b) Supongamos que los mensajes se codifican de manera que los errores provocados por un solo impulso de ruido se pueden corregir. Determinar la probabilidad de que un mensaje que dura t segundos sea recibido sin errores.
    c) Se consideran 100 intervalos consecutivos de longitud 1 segundo. ¿cuál es la probabilidad de que en mas de la mitad de ellos haya habido mas de un impulso ? Suponga que $\lambda = 2$ impulsos por segundo.

11. Un negocio tiene dos líneas telefónicas. Suponga que en un instante n puede haber 0, 1 ó 2 líneas ocupadas. Se observan las líneas ocupadas a intervalos regulares de 5 minutos. Sea X_n el número de líneas ocupadas en el instante n. Suponga que el proceso se puede describir como una cadena de Markov con probabilidades de transición estacionarias. Los estados posibles del proceso son 0, 1 y 2 y la matriz de transición de la cadena es:

$$P = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.2 & 0.8 & 0.0 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}.$$

a) Realice el diagrama de transición de estados del proceso.

b) ¿Cuál es la probabilidad de que en el instante $n = 1$ haya dos líneas ocupadas si al inicio del proceso había una línea ocupada?

c) Cuál es la probabilidad de que en el instante $n = 4$ haya una línea ocupada si en el instante $n = 3$ había dos líneas ocupadas?

d) ¿Cuál es la probabilidad de que en el instante $n = 4$ haya dos líneas ocupadas si en el instante $n = 2$ había dos líneas ocupadas?

e) Si $\pi(n)$ es la distribución de probabilidades de estados en el instante n, obtener $\pi(1)$, $\pi(2)$ y $\pi(3)$ si $\pi(0) = (0.3\ 0.3\ 0.4)$

f) Encuentre, si existe, un vector de probabilidades estacionario (la distribución de probabilidades de estados a *largo plazo*).

Nota: Si 1 es autovalor simple de la matriz de transición P entonces la distribución de probabilidades de estados a *largo plazo* es el autovector de la traspuesta de P cuyas componentes tienen sus componentes positivas y de suma 1. Para este ejemplo puede probar el siguiente código de *Octave*:

```
P = [0.5 0.3 0.2;0.2 0.8 0;0.3 0.3 0.4];
[a b] = eig(P');
a
b
pestac=a(:,1)/sum(a(:,1))
```

La matriz **b** que devuelve el procedimiento **eig** es diagonal y los elementos diagonales son los autovalores de P. En la primera columna de **a** se encuentran las componentes de un autovector de la traspuesta de P correspondiente al autovalor 1 de P. Dividiendo esas componentes por su suma se obtiene la distribución de probabilidades de estados a *largo plazo* para esta cadena de Markov.

12. Una empresa produce cierto artículo en cada período de ventas. Si logra un éxito de ventas en un período entonces la probabilidad de tener éxito en el próximo período es 0.5. Si en el período de ventas no tiene éxito entonces la probabilidad de tener éxito en el próximo período es 0.25. Supongamos que se considere que en el presente período se ha obtenido éxito.

    a) Describir el problema usando una cadena de Markov. Obtener la matriz P de transición y hacer un diagrama de transición de estados de la cadena.

    b) Obtener la probabilidad de *éxito* luego de tres períodos de ventas.

    c) Determinar con que frecuencia se tiene éxito de ventas *a tiempo grande.*

13. Tres supermercados S_1, S_2 y S_3 compiten por los clientes. Una investigación determina que al comenzar el mes de agosto los tres supermercados tienen igual cantidad de clientes. Al finalizar el mes se observa que:

a) S_1 conserva el 80 % de sus clientes y gana el 10 % y el 2 % de los clientes de S_2 y S_3 respectivamente.

b) S_2 conserva el 70 % de sus clientes y gana el 14 % y el 8 % de los clientes de S_1 y S_3 respectivamente.

c) S_3 conserva el 90 % de sus clientes y gana el 6 % y el 20 % de los clientes de S_1 y S_2 respectivamente.

Sea P es la matriz cuadrada de elementos p_{ij}, donde p_{ij} es la probabilidad de que un cliente del supermercado S_i se pase al supermercado S_j al cabo de un mes.

a) Construya la matriz de transición P, con los datos del problema.

b) Si $P^{(n)}$ es la matriz cuyo elemento de la posición i, j indica la proporción de clientes que se pasaron del supermercado S_i al S_j al cabo de n meses (bajo el supuesto de que estas proporciones permanecen invariables mes a mes), determine qué porcentaje de clientes se pasaron de S_2 al S_3 al cabo de 2 meses.

c) Sea $A = (1/3\ 1/3\ 1/3)$ el vector fila cuyos elementos indican la proporción de clientes que tenía inicialmente cada supermercado. El producto $AP^{(n)}$ indica la proporción de clientes de cada supermercado al cabo de n meses. Calcule qué proporción de clientes tiene cada supermercado al cabo de un año, dos años, tres años y cuatro años.

d) ¿Qué puede concluir acerca de la proporción de clientes de cada supermercado a largo plazo?

e) Verifique que la respuesta del punto d) también puede obtenerse resolviendo el sistema de ecuaciones $xP = 1.x$, donde $x = (x_1\ x_2\ x_3)$ y $x_1 + x_2 + x_3 = 1$, es decir, x es un autovector fila a izquierda de la matriz P correspondiente al autovalor 1.

14. Tres compañías, 1, 2 y 3, introducen al mercado simultáneamente marcas nuevas de pasta dental. Al principio, las proporciones iniciales del mercado en cada marca son de 40, 20 y 40 % respectivamente. Durante el primer año, la compañía 1 mantuvo el 85 % de su clientela, obtuvo el 15 % de la clientela de la compañía 2, y el 5 % de la compañía 3; la compañía 2 obtuvo el 5 % de la clientela de la compañía 1, retuvo el 75 % de su propia clientela, y obtuvo el 5 % de la de la compañía 3; y la compañía 3 obtuvo el 10 % de la clientela de la compañía 1, el 10 % de la compañía 2 y retuvo el 90 % de su propia clientela. Suponiendo que el mercado total que se comparte en estas tres compañías no varía y que cada año se intercambian entre ellas las mismas fracciones:

a) Encuentre la matriz de transición P.

b) Determine las proporciones del mercado en cada compañía después de: i) un año, ii) 2 años.

c) Determine las proporciones del mercado en cada compañía a largo plazo.

15. *El vendedor viajero* La región de ventas de un vendedor la componen tres ciudades A, B y C. Nunca vende en la misma ciudad en días seguidos. Si vende en la ciudad A, entonces al día siguiente vende en la ciudad B. Sin embargo, si vende en una de las dos ciudades B ó C, entonces al día siguiente la probabilidad de vender en A es el doble de la de vender en la restante ciudad.

a) Describir el problema usando una cadena de Markov. Obtener la matriz P de transición y hacer un diagrama de transición de estados de la cadena.

b) Verificar que existe una distribución de probabilidades de estados *a tiempo grande*.

c) Determinar con que frecuencia vende en cada ciudad *a tiempo grande*.

16. El estado j de una cadena de Markov $\{X_n\}$ se denomina *alcanzable* desde el estado i si es posible que se produzca una transición desde el estado i al estado j en un número finito de pasos. Para que esto ocurra se debe cumplir que el elemento i, j de la matriz de transición de n pasos $P^{(n)}$ sea positivo para algún $n > 0$. Si todo estado es alcanzable desde cualquier otro estado entonces se dice que la cadena es *regular*. ¿Es regular la cadena cuya matriz de transición se muestra a continuación?

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0.3 & 0 & 0.7 \\ 1 & 0 & 0 \end{bmatrix}.$$

17. Considere una cadena de Markov con espacio de estados $E = \{a,\ b,\ c\}$ y matriz de transición de probabilidades dada por

$$P = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 1 & 0 & 0 \\ 0 & 0.3 & 0.7 \end{bmatrix}.$$

a) Calcular $P(X_2 = a/X_1 = b, X_0 = c)$.

b) Calcular $P(X_{35} = a/X_{33} = a)$.

c) Estimar $P(X_{200} = a/X_0 = b)$.
(Probar primero que se alcanza un estado estacionario).

18. Una urna contiene 5 bolillas negras y 5 bolillas blancas. Consideremos el siguiente experimento,que se repite indefinidamente: se escoge aleatoriamente una bola de la urna; si es blanca, se vuelve a poner en la urna; si es negra, se deja fuera. Sea X_n el número de bolas negras en la urna después de n extracciones.

a) Describir el experimento como una cadena de Markov.

b) Determinar la matriz P de probabilidades de transición de estado.

c) Calcular $P_{54}^{(}2)$ (la probabilidad de que el proceso pase del valor 5 al valor 4 en dos etapas) de dos formas: directamente y calculando la matriz P^2.

d) Obtener la distribución de probabilidades de estados a *largo plazo*.

19. Jorge tiene 3 pesos. En cada jugada puede perder 1 peso con probabilidad 3/4 pero puede ganar dos pesos con probabilidad 1/4. Deja de jugar si pierde sus tres pesos o si gana por lo menos 3 pesos.

a) Obtener la matriz de transición de la cadena de Markov.
Pista: Considere 8 estados donde cada uno corresponde a que el jugador tenga 0, 1,..., 7 pesos. La matriz de transición tendrá muchos ceros.

b) ¿Cuál es la probabilidad de que haya por lo menos 4 jugadas en el juego?

20. Considere una cadena de Markov con espacio de estados $E = \{1,\ 2,\ 3\}$ y matriz de transición de probabilidades dada por

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}.$$

a) Usando solo el álgebra del cálculo de probabilidades y todas las posibles transiciones generar tablas para la probabilidad de la transición $i \to j$ en n pasos para los siguientes pares (i, j): $(2, 1)$, $(3, 3)$, $(1, 1)$, y $(2, 2)$ y para n tomando los valores 1 a 5.

b) Probar que la cadena es regular y obtener la distribución de probabilidades de estados a *largo plazo*.

21. La matriz de probabilidades de transición de un paso de una cadena de Markov es:

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 1 & 0 \end{bmatrix}.$$

a) Demostrar que la cadena es regular.

b) Si el proceso comienza en el estado 1 determinar la probabilidad de alcanzar el estado 3 en dos pasos.

c) Obtener la distribución de probabilidades a *largo plazo*.

22. Considere una cadena de Markov con espacio de estados $E = \{0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\}$ y matriz de transición de probabilidades dada por

$$P = \begin{bmatrix} \frac{1}{5} & \frac{3}{5} & 0 & 0 & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

a) Represente el diagrama de transición de estados de la cadena.

b) Por inspección del diagrama verifique que:

1) el conjunto de estados tiene tres subconjuntos disjuntos (se denominan *clases*) que corresponden a estados comunicados entre ellos.

2) Hay dos clases que son cerradas y recurrentes, una vez que uno de los estados de la clase cerrada es alcanzado entonces todos los estados subsiguientes de la cadena corresponden a esa clase.

3) La clase restante es transiente, el proceso puede terminar saliendo de ella.
4) Una de las clases cerradas es periódica de periodo 3, la otra no es periódica.

c) Usando solo el álgebra del cálculo de probabilidades y la suma de alguna serie numérica demuestre que si el proceso comienza en:

1) el estado 0 entonces la probabilidad de alcanzar el estado 6 es $\frac{1}{4}$.
2) el estado 1 entonces la probabilidad de alcanzar el estado 3 es 1.
3) el estado 1 entonces el numero promedio de pasos hasta llegar al estado 3 es 3.
4) el estado 1 entonces a largo plazo la probabilidad de que el proceso se encuentre en el estado 2 es $\frac{3}{8}$.

Son datos

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \qquad |q| < 1$$
$$\sum_{k=1}^{\infty} k\, q^k = \frac{q}{(1-q)^2} \qquad |q| < 1$$

23. Un aficionado utiliza un método bastante simple para pronosticar el tiempo atmosférico. Clasifica cada día como *seco* o *húmedo*, y supone que la probabilidad de que cualquier día dado sea igual al precedente está dada por una constante p, $(0 < p < 1)$. De acuerdo con registros anteriores, se sabe que la probabilidad de que el primero de enero sea *seco* es β $(0 < \beta < 1)$.

a) Si $\beta_n = P($ el n-ésimo día del año sea *seco*$)$, obtener una expresión para β_n en función de p y n, para $n = 1, 2, 3$.

b) Si el tercer día del año fue *seco*, ¿cuál es la probabilidad de que el primero de enero también haya sido *seco*?

c) Analizar la situación como una cadena de Markov y obtener la distribución de probabilidades estacionaria.

24. Suponga este proceso de dos estados: M, *muerto* y V, *vivo*. Estos son los estados de una persona cuando comienza un año. Sea entonces el proceso estocástico donde X_n es el estado al comienzo del año n y que en $n = 0$ el estado es V. Se considera que el proceso es de Markov y que las probabilidades de transición de un paso no nulas son: $P(X_n = V/X_{n-1} = V) = 1 - p$, $P(X_n = M/X_{n-1} = V) = p$ y que $P(X_n = M/X_{n-1} = M) = 1$ dado que el estado M es claramente absorbente.

a) Obtener P^n con P la matriz de probabilidades de transición de un paso. Interpretar su valor límite. Para obtener P^n se sugiere calcular unas primeras potencias, luego conjeturar la forma y demostrar usando inducción completa.

b) Obtener una expresión para $p_n = P(X_{n+1} = M \cap \bigcap_{k=1}^{n} X_k = V/X_0 = V)$.
Sugerencia: $p_1 = P(X_2 = M \cap X_1 = V/X_0 = V) = P(X_2 = M/X_1 = V)\, P(X_1 = V/X_0 = V) = (1-p)\, p$. Esta es la probabilidad de la persona se muera antes de cumplir 2 años.

c) Sea N la variable aleatoria que cuenta los años de vida de una persona. Probar que $P(N = 0) = p$, y que para $k \in \mathbf{N}$ se tiene $P(N = k) = p\,(1-p)^k$. Calcular el valor esperado de N.

25. Cada materia que cursa un alumno en una universidad tiene tres oportunidades para dar el examen final. Suponga que la probabilidad de aprobar el examen final es siempre p. Sea X_n la variable aleatoria que da el número de oportunidades que tiene el alumno en el período n. El recorrido de X_n es el conjunto $\{0, 1, 2, 3\}$ siendo el valor cero el estado que se alcanza cuando se aprueba el examen final (claramente un estado absorbente). El estado 3 corresponde al que se tiene una vez aprobada la cursada. Cuando no se aprueba en la última de las instancias se produce una transición del estado 1 al 3 (la materia se recursa).

    a) Modelar la evolución de este proceso como una cadena de Markov obteniendo la matriz de probabilidades de transición de un paso.

    b) Suponga que el estado inicial es el 3. En este caso la distribución de probabilidades es $(0\,,0\,,0\,,1)$. Obtener la distribución de probabilidades para los primeros tres períodos y conjeturar sobre su forma para todo n. Analizar su valor límite.