

Final probabilidad y estadística

Martín Rossi

Parte I

1.

i.

$$P = \begin{pmatrix} q & p & 0 & 0 & \dots \\ q & 0 & p & 0 & \dots \\ q & 0 & 0 & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix}$$

Donde $E = \{0, 1, 2, \dots\}$

Para calcular N_k primero calcularía P^k para obtener la matriz de transición en k pasos, y me fijaría en la primera fila la columna k , que sería la probabilidad de llegar desde 0 hasta k éxitos consecutivos. Por lo tanto si hago 1 sobre esa probabilidad me daría $E(N_k)$.

ii.

$$P = \begin{pmatrix} q & p & 0 & 0 & \dots \\ 0 & q & p & 0 & \dots \\ 0 & 0 & q & p & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix}$$

Es decir que en caso de no éxito no se vuelve a 0 sino que queda en la misma cantidad.

iii.

La distribución en ese caso sería la pascal, por lo que la esperanza es $E(M_k) = k/p$, donde se usa M_k para diferenciarla de N_k , y p es la probabilidad de éxito.

iv.

La esperanza de M_k será menor o igual a la de N_k , al no tener la restricción de ser éxitos consecutivos.

2.

i.

El proceso S_n muestra la posición de la partícula al momento n , que se basa en el proceso de bernoulli I_n , o sea que la relación que tiene con el número de éxitos es que será la cantidad de veces que la partícula se movió en una dirección en la recta, es decir, que si el número de éxitos fue k la partícula se habrá movido k veces en la misma dirección. Pero no determinará su posición.

ii.

En el trabajo se usó $p = 0.85$

S_n se definió como $S_n = D_0 + \dots + D_n$ por lo tanto por las propiedades de la esperanza:

$$E(S_n) = E(D_0 + \dots + D_n) = E(D_0) + \dots + E(D_n) = E(2I_0 - 1) + \dots + E(2I_n - 1) = 2E(I_0) - 1 + \dots + 2E(I_n) - 1 = n * 2 * 0.85 - n = 0.7n$$

A a largo plazo se esperaría ver que S_n tienda a infinito porque hay más probabilidades que se mueva hacia adelante que hacia atrás.

3.

i.

$$R_A = \cup_{i=0}^{\infty} X_i \leq 0$$

ii.

Para S el objetivo del jugador

$$E = \{0..S\}$$

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & q & 0 & p & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

iii.

Lo que se busca calcular es la $F(i, j)$. Para ello se busca primero la matriz R , de visitas promedio entre estados y después se calcula $F(i, j) = R(i, j)/R(i, i)$. Y aproximaría la matriz R con $R = \sum_{m=0}^{\infty} P^m(i, j)$ ya que no se puede con $R = (I - P)^{-1}$ porque no existe esa inversa.

iv.

Para calcular la cantidad promedio de pasos que tomaría llegar a un estado absorbente se calcula para cada estado inicial i : $\sum_{n=0}^{\infty} n * (F_n(i, 0) + F_n(i, S))$ y se construye la matriz, que tiene sólo una fila. Lo que se está calculando es la suma de probabilidades de llegar a 0 o S por primera vez en n pasos, que son los estados absorbentes (esto se puede hacer porque son excluyentes), y se multiplica por n , que serían los pasos hasta llegar a alguno de ellos, obteniendo así la esperanza.

4.

Las dos variables serían M_n : la cantidad de comparaciones para una lista de tamaño n , y P_n : la posición del pivote para una lista de tamaño n .

5.

i.

Que para simular los saltos entre las páginas se usaba siempre la misma matriz de transición P , y no se tenía que andar calculando cada vez la matriz. Además es más fácil comparar entre simulaciones ya que la matriz no varía de la trayectoria recorrida hasta un momento dado de la simulación.

ii.

Sí, la distribución tiene límite porque la cadena es ergódica. La distribución se calculó en el ejercicio 5.c del trabajo, por lo que la matriz P^{∞} sería la matriz 7×7 con todas las filas iguales a la π^{∞} .

6.

i.

El conjunto de momentos de observación es continuo (tiempo).

El conjunto de los valores que puede tomar el proceso es discreto (cantidad de aterrizajes).

ii.

La distribución que se ajusta al tiempo entre llegadas es la exponencial, como se puede ver en el gráfico del ejercicio del trabajo.

iii.

En el momento inicial de observación hay 0 aterrizajes.

La cantidad de aterrizajes es creciente en el tiempo, es decir, $s \leq t \implies N_s \leq N_t$.

Los tiempos entre aterrizajes son variables aleatorias independientes y tienen la misma distribución(exponencial).

No se pueden producir dos arribos al mismo momento.

7.

i.

Los estados son todos recurrentes.

ii.

Los estados de la cadena forman un conjunto cerrado e irreducible.

Parte II

a.

Para calcular la altura de las barras en un histograma se debe hacer $\frac{\text{frecuencia relativa}}{\text{ancho de la clase}}$, lo que se debería hacer para cada clase elegida.

Y en el caso particular de la clase $[65, 80]$ sería $\frac{\text{frecuencia relativa de usuarios entre 65 y 80}}{15}$

b.

Se podría hacer un gráfico de dispersión donde se muestren los elementos de cada sexo agrupados, pero separados del otro. Así se podrían comparar viendo la totalidad de la población y sin tener que categorizarlos.