Procesos de Bernoulli

Considere un experimento que consiste:

✓una sucesión infinita de ensayos realizados en idénticas condiciones.
✓los ensayos son independientes entre sí y
✓sólo pueden tener dos resultados posibles:

"éxito" (E) o "fracaso"(F)

con: $P(E) = p$, $P(F) = 1 - p$,
constante a lo largo de todo el proceso.

Este experimento puede modelarse como la sucesión de infinitos ensayos de Bernoulli.

El espacio muestral común de este experimento es

$$\Omega = \{w : w = (w_1, w_2, ..., w_n, ...),\ w_i = E \text{ o bien } w_i = F, i \in \mathrm{N}\}.$$

A la n-ésima repetición del ensayo podemos asociarle una variable aleatoria X_n que toma dos valores posibles

$$X_n(w) = \begin{cases} 1, & \text{si } w_n = E, \\ 0, & \text{si } w_n = F. \end{cases} \quad [1]$$

$$P(X_n = 1) = p;\ P(X_n = 0) = 1 - p. \quad [2]$$

Se denomina *proceso de Bernoulli* a una colección numerable de variables aleatorias independientes, definidas como en [1] con distribución de prob. [2].

Ejemplo

En una bifurcación de una ruta, aproximadamente el 62% de los automóviles toma la rama izquierda.

Se define:

$X_n = 1$ si el n-ésimo automóvil toma la rama izquierda,

$X_n = 0$ si toma la rama derecha.

Se supone que los conductores eligen su camino independientemente de lo que hacen los otros, eso quiere decir que se puede considerar

$X_1, X_2, \ldots, X_n, \ldots\ldots$ independientes con $P(X_n = 1) = 0.62 \quad \forall n$

El proceso $\{X_n : \quad n \in \mathbb{N}\}$ es un *proceso de Bernoulli.*

Proceso Número de Éxitos

Sea $\{X_n : n \in \mathbb{N}\}$ un proceso de Bernoulli con P(E) = p.

La variable aleatoria:

$$N_n = X_1 + X_2 + \ldots + X_n \qquad [3]$$

describe el número de éxitos en los primeros n ensayos de Bernoulli.
Si además se define $N_0 = 0$, el proceso estocástico

$$\{N_n : n \in \mathrm{N}_0\}$$

describe el número de éxitos en un proceso de Bernoulli y tiene a $\mathbb{N}_0$ como espacio de estados.

Notar que (3) indica que:

$$N_n = k \Leftrightarrow X_1 + X_2 + \ldots + X_n = k \Leftrightarrow$$

exactamente k de los n sumandos es igual a 1 y, por lo tanto, se deduce:

$$N_n \sim Bi(n, p)$$

$$P(N_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \ \forall k = 0, \ldots, n.$$

Las variables aleatorias

$$N_{m+n} - N_m$$

representan el número de éxitos entre el m-ésimo y el (m + n)-ésimo ensayo. Por lo tanto:

$$P(N_{m+n} - N_m = k) = \binom{n}{k} p^k (1-p)^{n-k}, \ \forall k = 0, ..., n,$$

lo cual significa que el número de éxitos sólo depende de la cantidad de ensayos observados y no del instante en que comenzó a observarse el proceso.

Para pensar!

¿Influye la cantidad de éxitos ocurridos en los m primeros ensayos sobre la cantidad de los que ocurrirán entre la m*-ésima* y la (m+n) *-ésima* repetición del ensayo de Bernoulli?

La independencia de las Xi asegura que no.

$$P(N_{m+n}-N_m=k|N_0,N_1,...,N_m)=P(N_{m+n}-N_m=k),\ \forall k=0,...,n.$$

dado que el valor de la probabilidad es independiente de m se dice que el proceso es a ***incrementos estacionarios***

Más aún, las variables aleatorias:

$$N_{n_1}, N_{n_2} - N_{n_1}, \ldots, N_{n_m} - N_{n_{m-1}} \qquad (0 < n_1 < n_2 < \ldots < n_m)$$

son variables aleatorias ***independientes***, por lo tanto:

$$P(N_{m+1} = k | N_0, N_1, \ldots, N_m) = P(N_{m+1} = k | N_m).$$

esta igualdad indica que el futuro inmediato del proceso depende sólo del presente y no del pasado.

⇩

propiedad markoviana

Si se conoce cuántos éxitos hubo hasta el instante n-*ésimo,* ¿cuántos éxitos podemos tener al instante siguiente?

$$P(N_{n+1}=j|N_n=i)=P(N_n+X_{n+1}=j|N_n=i)=P(X_{n+1}=j-i)$$

$$P(X_{n+1}=j-i)=\begin{cases} p & \text{si } j-i=1 \\ 1-p & \text{si } j-i=0 \\ 0 & \text{en otro caso,} \end{cases}$$

$$P(N_{n+1}=j|N_n=i)=\begin{cases} p & \text{si } j=i+1 \\ 1-p & \text{si } j=i \\ 0 & \text{en otro caso} \end{cases}$$

Notar que estas probabilidades son independientes de n.

- ✓ Estas probabilidades, que pueden interpretarse como la ***probabilidad de transición*** en un paso entre dos estados del proceso.
- ✓ Estas probabilidades pueden representarse en una matriz o a través de un grafo.

Propuesta:

- Armar la matriz de transición

$$P(i,j) = P(N_{n+1} = j | N_n = i)$$

- Construya el grafo correspondiente al proceso planteado

El promedio de N_{n+1} dado que se conoce N_n es:

$$
\begin{aligned}
E(N_{n+1}|N_n) &= E(N_n + X_{n+1}|N_n) = E(N_n|N_n) + E(X_{n+1}|N_n) \\
&= N_n + p.
\end{aligned}
$$

En forma análoga:

$$E(N_{n+m}|N_n) = E(N_n|N_n) + E(N_{n+m} - N_n|N_n) = N_n + mp.$$

$$E(N_{n+m}|N_0, N_1, ..., N_n) = E(N_{n+m}|N_n).$$

Instantes de Éxito

Sea $\{X_n : n \in \mathrm{N}\}$ un proceso Bernoulli con $P(E) = p.$
Sea la variable aleatoria
T_k : instante en que se produce el k-*esimo* éxito.
Para dicha variable aleatoria el espacio de estados es □ .
Es decir:

$$T_k = \min\{n \in \mathrm{N} : N_n = k\}$$

El proceso $\{T_k : k \in \mathrm{N}\}$ es el proceso que describe los instantes en los que se producen los éxitos de Bernoulli y se conoce como *instante de éxito*

Lemas:

Lema 1 *Si $n \geq k$, resulta*

$$T_k = n \Leftrightarrow N_{n-1} = k-1, X_n = 1,$$

$$T_k \leq n \Leftrightarrow N_n \geq k.$$

Lema 2 *Si $n \geq k$,resulta*

$$P(T_k = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k},$$

$$P(T_k \leq n) = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}.$$

Propuesta: A partir del Lema 1 demuestre el Lema 2

Instantes de Éxitos (cont.)

Por lo tanto, las variables aleatorias T_k tienen distribución de Pascal ya que representan el número de repeticiones hasta que aparece el suceso exitoso por k–*ésima* vez.

Para pensar:

Si se conoce los instantes en que se han producido los primeros k éxitos.

a. ¿Ayuda ese conocimiento para determinar la probabilidad de que el tiempo transcurrido entre el k-*ésimo* y el $(k+1)$-*ésimo* éxito tenga determinada longitud?
b. ¿Ayuda ese conocimiento para determinar la probabilidad de que el $(k+1)$-*ésimo* éxito se produzca en determinado instante?

$$
\begin{aligned}
P(T_{k+1} - T_k &= m \mid T_1, T_2, ..., T_k) = P(T_{k+1} - T_k = m) \\
P(T_{k+1} - T_k &= m) = p(1-p)^{m-1} \\
P(T_{k+1} &= n \mid T_1, T_2, ..., T_k) = P(T_{k+1} = n \mid T_k) \\
P(T_{k+1} &= n \mid T_k) = \begin{cases} 0 & T_k \geq n \\ p(1-p)^{n-1-T_k} & T_k < n \end{cases}
\end{aligned}
$$

Instantes de Éxitos (cont.)

Se puede deducir de las igualdades anteriores que el proceso tiempos de éxito goza de la propiedad Markoviana y

$$P(T_{k+1} = j \mid T_k = i) = \begin{cases} 0 & i \geq j \\ p(1-p)^{j-i-1} & i < j \end{cases}$$

indica que las probabilidades de transición en un paso son independientes del número de éxito del que se está estudiando el instante de ocurrencia, independiente de k.

Instantes de Éxitos (cont.)

Además el proceso tiene incrementos independientes y que los incrementos tienen *distribución geométrica*.
En cuanto a esperanzas condicionadas, si $m > k$

$$
\begin{aligned}
E(T_m|T_1, T_2, ..., T_k) &= E(T_m|T_k) = E((T_m - T_k) + T_k|T_k) \\
&= E((T_m - T_k)|T_k) + E(T_k|T_k) \\
&= \frac{m-k}{p} + T_k.
\end{aligned}
$$

Suma de variables aleatorias independientes

Los procesos antes estudiados tienen ciertas características en común:
Las variables aleatorias que los componen pueden ser presentadas como sumas de otras variables aleatorias iid.

$$
\begin{aligned}
N_n &= N_0 + (N_1 - N_0) + (N_2 - N_1) + \ldots + (N_n - N_{n-1}) \\
&= X_1 + X_2 + \ldots + X_n, \\
T_k &= T_1 + (T_2 - T_1) + (T_3 - T_2) + \ldots + (T_k - T_{k-1})
\end{aligned}
$$

Generalización

Si se considera una sucesión de variables aleatorias:

$$\{Y_n : n \in \mathbb{N}\}$$

iid, y para cada $n \in \mathbb{N}$ se define:

$$Z_n = \begin{cases} 0 & n = 0 \\ Y_1 + Y_2 + ... + Y_n & n \geq 1. \end{cases}$$

El proceso $\{Z_n : n \in \mathbb{N}\}$ goza de las siguientes propiedades:

Lema 3 *El proceso $\{Z_n : n \in \mathbb{N}\}$ tiene incrementos independientes y estacionarios.*

Lema 4 *Si $W = g(Z_n, Z_{n+1}, ...)$, entonces $E(W|Z_0, Z_1, ..., Z_n) = E(W|Z_n)$*