PROCESO DE BERNOULLI
Rosario Romera
Febrero 2009

1. Sumas de Variables Aleatorias Independientes

Definición Se considera el experimento aleatorio consistente en la repetición de juegos binarios independientes. El resultado de cada realización será un elemento de Ω siendo

$$\Omega \quad = \{\omega, \omega = (\omega_1, \omega_2 \ldots) \quad \text{donde} \quad \omega_i \text{ es } A \quad \text{ó} \quad \overline{A}\}$$

Para cada subconjunto de Ω (suceso) definimos la siguiente medida de probabilidad

$$Prob(\omega_1 \quad \omega_2 \ldots) \quad = \quad Prob(\omega_1).Prob(\omega_2)\ldots$$

siendo: $\quad Prob(A) = p \qquad Prob(\overline{A}) = 1 - p = q$

Por ejemplo: $\quad Prob(A\overline{A}AA) = p^3 q$

donde se consideran las variables aleatorias independientes Bernoulli $\quad (p)$:

$$X_i(A) = 1, \qquad X_i(\overline{A}) = 0$$

con $\quad Prob(X_i = 1) = p, \qquad Prob(X_i = 0) = 1 - p \qquad \forall i \in \mathbb{N}$

Definición Proceso de Bernoulli. La sucesión aleatoria $\{X_n\}_{n\geq 1}$ es Proceso Bernoulli con probabilidad de éxito p, si verifica:

1. $X_1, X_2, \ldots$ son independientes e idénticamente distribuidas (iid).
2. $P(X_n = 1) = p, \qquad P(X_n = 0) = q = 1 - p \qquad \forall n$

Propiedades.

1. $E[X_n^r] = p \qquad \forall n \geq 1 \quad y \quad \forall r \geq 1$
2. $Var[X_n] = E[X_n^2] - E^2[X_n] = p - p^2 = p.q$
3. $G_{Xn}(s) = E[S^{Xn}] = sp + q \qquad \forall s \geq 0$

Observación: es proceso estrictamente estacionario y

$$cov(X_n, X_s) = E[X_n X_s] - E[X_n]E[X_s] = 0$$

Proceso Sumas de Bernoulli Modeliza el número de éxitos en n juegos de Bernoulli independientes:

Definición:

$$\begin{aligned} S_n(\omega) &= X_1(\omega) + \ldots + X_n(\omega) \qquad n \geq 1, \qquad y \quad S_0(\omega) = 0 \\ S_{n+m} - S_n &= X_{n+1} + X_{n+2} + \ldots + X_{n+m} \\ &= \text{número de éxitos en los juegos } n+1, n+2, \ldots n+m \end{aligned}$$

Observación. El proceso $\{S_n\}_{n \in \mathbb{N}}$ es una cadena de parámetro discreto.

Propiedades

1. $E[S_n] = np$
2. $Var[S_n] = npq$
3. $G_{S_n}(s) = \prod_{i=1}^{n} G_X(s) = G_X^n(s) = (sp+q)^n$

Ejemplo de cálculo de momentos:

$$\begin{aligned} G'_{S_n}(1) &= [n(sp+q)^{n-1}.p]_{s=1} = np(p+q)^{n-1} = np \\ G''_{Sn}(1) &= [n(n-1)(sp+q)^{n-2}.p^2]_{s=1} \\ &= n(n-1)p^2(p+q)^{n-2} \\ &= n(n-1)p^2 \end{aligned}$$

entonces

$$\begin{aligned} E[S_n] &= G'_{S_n}(1) = np \\ Var[S_n] &= E[X^2] - E^2[X] \\ &= G''_x(1) + G'(1) - [G'^2 \\ &= n(n-1)p^2 + np - (np)^2 \\ &= n^2p^2 - np^2 + np - n^2p^2 \\ &= np(1-p) \\ &= npq \end{aligned}$$

Proposición:

$$\forall n, k \in \mathbb{N} \qquad P(S_{n+1} = k) = p.P(S_n = k-1) + qP(S_n = k)$$

Demostración: Por el Teorema de Probabilidad Total.

$$\begin{aligned}
P(S_{n+1}=k) &= \sum_{j\in N} P(S_{n+1}=k/S_n=j)P(S_n=j) \\
&= \sum_{j\in N} P(X_{n+1}=k-j)P(S_n=j)
\end{aligned}$$

$$\text{ya que:} \qquad P(X_{n+1}=k-j)=\begin{cases} p \text{ si } k=j+1 \\ \text{q si } k=j \\ 0 \text{ resto} \end{cases}$$

$$\begin{aligned}
&= P(X_{n+1}=1).P(S_n=k-1)+P(X_{n+1}=0).P(S_n=k) \\
&= p.P(S_n=k-1)+qP(S_n=k) \qquad \square
\end{aligned}$$

Proposición

$$\forall n\in\mathbb{N} \qquad P(S_n=k)=\frac{n!}{k!(n-k)!}p^kq^{n-k} \qquad \forall k=0,\ldots,n$$

Demostración (por inducción completa)

	$k=0$	$k=1$	$k=2$	$k=3$
$n=0$	1	0	0	0
$n=1$	q	p	0	0
$n=2$	q^2	$2qp$	p^2	0

supongamos que es cierta la fórmula para m, veamos que es cierta para $n=m+1$:

$$P(S_{m+1}=0)=q^{m+1}$$

para $0<k\leq m+1$, por la Proposición anterior, se tiene:

$$\begin{aligned}
P(S_{m+1}=k) &= p.P(S_m=k-1)+qP(S_m=k) \\
&= p\frac{m!}{(k-1)!(m-k+1)!}p^{k-1}q^{m-k+1}+q\frac{m!}{k!(m-k)!}p^kq^{m-k} \\
&= \frac{m!}{(k-1)!(m-k+1)!}p^kq^{m-k+1}+\frac{m!}{k!(m-k)!}p^kq^{m-k+1} \\
&= \frac{m!}{(k-1)!(m-k)!}p^kq^{m-k+1}\left[\frac{1}{m-k+1}+\frac{1}{k}\right] \\
&= \frac{m!(m+1)}{(k-1)!k(m-k)!(m-k+1)}p^kq^{m-k+1} \\
&= \frac{(m+1)!}{k!(m-k+1)!}p^kq^{m-k+1} \qquad \square
\end{aligned}$$

Corolario

$$\forall n,m\in\mathbb{N} \qquad P(S_{n+m}-S_m=k)=\binom{n}{k}p^kq^{n-k}, \qquad \forall k=0\ldots n$$

(ya que $S_{n+m}-S_m=X_{m+1}+X_{m+2}+\ldots+X_{m+n}\equiv$ suma de n variables aleatorias Bern(p) independientes)

Ejemplo.

$$P(S_{16}-S_9=3)=\left(\begin{array}{c}7\\3\end{array}\right)p^3q^4=35p^3q^4$$

Para hallar probabilidades conjuntas del tipo:

$$P(S_{n_1}=s_1,S_{n_2}=s_2,S_{n_3}=s_3)$$

establecemos el siguiente resultado

Proposición

$$\begin{array}{rcl}\forall n,m\in\mathbb{N} & & \\ P(S_{m+n}-S_m=k/S_0,\ldots S_m) & = & P(S_{m+n}-S_m=k)\\ & = & \left(\begin{array}{c}n\\k\end{array}\right)p^kq^{n-k},\qquad \forall k=0,\ldots,n\end{array}$$

Demostración

$S_0,\ldots,S_m$, están completamente determinadas por $X_1,\ldots,X_m$ y análogamente $X_1,\ldots,X_m$ están determinados por $S_0,\ldots,S_m$ (generan la misma σ-álgebra de sucesos), entonces:

$$\begin{array}{rcl}P(S_{m+n}-S_m=k/S_0\ldots S_m) & = & P(S_{m+n}-S_m=k/X_1\ldots X_m)\\ & = & P(X_{m+1}+X_{m+2}+\ldots+X_{m+n}=k/X_1\ldots X_m)\\ & = & (X_i\text{ son variables aleatorias independientes})\\ & = & P(X_{m+1}+X_{m+2}+X_{m+n}=k)\\ & = & P(S_{m+n}-S_m=k)\end{array}$$

Corolario Sean $0<n_1<n_2<\ldots<n_j$ pertenecientes a $\mathbb{N}$, con j arbitrario, entonces las variables

$$S_{n_1}-S_0,S_{n_2}-S_{n_1},...S_{n_j}-S_{n_{j-1}}$$

son independientes

Ejemplos

$$\begin{array}{rcl}P(S_{10}=4,S_{12}=5,S_{18}=8) & = & P(S_{10}=4,S_{12}-S_{10}=1,S_{18}-S_{12}=3)\\ & = & P(S_{10}=4).P(S_{12}-S_{10}=1).P(S_{18}-S_{12}=3)\\ & = & \left(\begin{array}{c}10\\4\end{array}\right)p^4q^6\quad\left(\begin{array}{c}2\\1\end{array}\right)pq\left(\begin{array}{c}6\\3\end{array}\right)p^3q^3\\ & = & (210,2,20)p^8q^{10}=8400p^8q^{10}\end{array}$$

$$\begin{array}{rcl}E[S_5S_8] & = & E[S_5(S_5+(S_8-S_5))]\\ & = & E[S_5^2+S_5(S_8-S_5)]\\ & = & (\text{por independencia})\\ & = & E[s_5^2]+E[S_5].E[S_8-S_5]\\ & = & \ldots=40p^2+5qp\end{array}$$

Definición El Proceso Estocástico $\{Z_n\}_{n\in N}$, es de **incrementos independientes** si:

$$0 < n_1 < n_2 < \ldots < n_j, \quad j \in \mathbb{N} \quad Z_{n_1} - Z_0, Z_{n_2} - Z_{n_1}, ..., Z_{n_j} - Z_{n_{j-1}}$$

son variables aleatorias independientes.

Definición El Proceso Estocástico $\{Z_n\}_{n\in N}$, es de **incrementos estacionarios** si:

$$\forall h \in \mathbb{N}, \quad \forall 0 < n_1 < n_2 < \ldots < n_j, \quad j \in \mathbb{N}$$

la distribución de

$$Z_h - Z_0, Z_{n_1+h} - Z_{n_1}, Z_{n_2+h} - Z_{n_2} ..., Z_{n_j+h} - Z_{n_j}$$

sólo depende de h y no de n_j.

Proposición El Proceso S_n (suma de variables aleatorias Bernoulli) es de Incrementos Independientes y Estacionarios.

Ejercicio Demostrar que si $\{Z_n\}_{n\in\mathbb{N}}$ tiene incrementos independientes y estacionarios y además $E[Z_n] < \infty \quad \forall n$, entonces:

$$\begin{aligned} E[Z_n] &= E[Z_0] + m_1.n, \quad \text{donde } m_1 = E[Z_1] - E[Z_0] \\ Var[Z_n] &= Var[Z_0] + \sigma_1^2 n \quad \text{donde } \sigma_1^2 = E[(Z_1 - m_1)^2] - Var[Z_0] \end{aligned}$$

Ejercicio Si $\{X_i\}$ son variables aleatorias iid con valores en $\mathbb{N}$, y $Z_n = X_1 + X_2 + \ldots + X_n$, con $n \in \mathbb{N}$ entonces Z_n es de incrementos independientes y estacionarios.

Otro punto de vista sobre el Proceso de Bernoulli: interés en predecir el futuro del proceso de sumas, según la información del pasado.

Proposición: (Propiedad de Markov) Sea $Y = g(S_m, S_{m+1}, \ldots, S_{m+n})$ para algún $n \in \mathbb{N}$, esto es, Y es variable aleatoria que depende de un número finito de variables. Entonces $E[Y/S_0 \ldots S_m] = E[Y/S_m]$

Demostración

Por el argumento de identidad de las σ-álgebras

$$\sigma \sim (X_1 \ldots X_m) \quad \equiv \quad \sigma \quad \sim (S_0 \ldots S_m)$$

se verificará que para cada función g existirá una cierta función g^* tal que:

$$g(S_m, S_{m+1} \ldots, S_{m+n}) = g^*(S_m, X_{m+1}, \ldots, X_{m+n})$$

luego para cada valor de Y se tiene que:

$$\begin{aligned} P(S_m &= k, X_{m+1} = i_1, \ldots X_{m+n} = i_n / S_0 \ldots S_m) \\ &= \text{(por ser } \{Xi\} \text{ independientes)} \\ &= P(S_m = k/S_0 \ldots S_m).P(X_{m+1} = i_1) \ldots P(X_{m+n} = i_n) \end{aligned}$$

luego

$$E[Y/S_0 \ldots S_m] = \text{función de } (S_m) = E[Y/S_m] \qquad \square$$

Ejemplos: Calcular

$$\begin{aligned}
E[S_{11}/S_3S_4S_5] &= \text{(por la propiedad markoviana)} = E[S_{11}/S_5] \\
&= E[S_5 + (S_{11} - S_5)S_5] = E[S_5/S_5] + E[S_{11} - S_5/S_5] \\
&= S_5 + E[S_{11} - S_5] = S_5 + 6p
\end{aligned}$$

$$\begin{aligned}
E[S_5S_8] &= E[E[S_5S_8/S_5]] = E[S_5.E[S_8/S_5]] \\
&= E[S_5(S_5 + 3p)] = E[S_5^2] + 3pE[S_5] \\
&= Var[Bin(5,p)] + E^2[Bin(5,p)] + 3p{,}5p = 40p^2 + 5pq
\end{aligned}$$

2. Proceso del número de intentos hasta obtener k éxitos sucesivos en juegos independientes Bernoulli

Sea $T_k(\omega)$ = número de intentos hasta el k-ésimo éxito, por ejemplo, si se ha obtenido en una experimentación

$$X_1(\omega) = 0, \quad X_2(\omega) = 1, \quad X_3(\omega) = 0, \quad X_4(\omega) = 1 \ldots$$

entonces $T_1(\omega) = 2, \quad T_2(\omega) = 4, \ldots, \{T_k\}_{k \in \mathbb{N}}$ contabiliza la cantidad de k éxitos.

Lema $\forall \omega \in \Omega$, con $k \geq 1$ y $n \geq k$ se verifica que:

$$\begin{array}{llcl} (1) & T_k(\omega) \quad \leq n & \Longleftrightarrow & S_n(\omega) \geq k \\ (2) & T_k(\omega) = n & \Longleftrightarrow & S_{n-1}(\omega) = k-1, X_n(\omega) = 1 \end{array}$$

Teorema

$$\begin{aligned} & \forall k \geq 1 : \\ P\{T_k \quad \leq n\} &= \sum_{j=k}^{n} \binom{n}{j} p^j q^{n-j}, \qquad n = k, k+1, \ldots \\ P\{T_k = n\} &= \binom{n-1}{k-1} p^k q^{n-k}, \qquad n = k, k+1, \ldots \end{aligned}$$

Distribución Binomial Negativa.

Demostración. Fijando $k \in \{1, 2, \ldots\}$ y $n \geq k$ entonces:

$$P\{T_k \leq n\} = P\{S_n \quad \geq k\} = \sum_{j=k}^{n} P\{S_n = j\}$$

de donde se obtiene (1).

Como los sucesos $\{T_k(\omega) = n\}$ y $\{S_{n-1} = k-1, X_n = 1\}$ son idénticos, se deduce que:

$$\begin{aligned} Prob\{T_k = n\} &\equiv Prob\{S_{n-1} = k-1, X_n = 1\} \\ &= P\{S_{n-1} = k-1\} P(X_n = 1) \\ &= \binom{n-1}{k-1} p^{k-1} q^{n-k}.P \qquad \square \end{aligned}$$

Observación

$$Prob(\text{el lugar que ocupa el } k\text{-ésimo} = n) = \binom{k+(n-k)-1}{n-k} p^k q^{n-k}$$

$\equiv Prob(n^o\text{fracasos previo } k\text{-éxito} = n-k)$

$\equiv Prob(n^o\text{fracasos antes del } k\text{-ésimo éxito} = x = n-k = 0, 1, \ldots)$

se verifica además que:

$$\binom{n-1}{k-1} = \binom{k+(n-k)-1}{k-1} = \binom{k+(n-k)-1}{n-k}$$

$$\begin{aligned}
G_{BN(k,p)}(s) &= E[S^x] = \sum_{n=k}^{\infty} s^n \begin{pmatrix} k+(n-k)-1 \\ n-k \end{pmatrix} p^k q^{n-k} \\
&= s^k \sum_{x=n-k=0}^{\infty} s^x \begin{pmatrix} k+x-1 \\ x \end{pmatrix} p^k q^x = p^k s^k \sum_{x=0}^{\infty} (sq)^x \begin{pmatrix} k+x-1 \\ x \end{pmatrix} \\
&= p^k (1-sq)^{-k} = \left(\frac{ps}{1-sq} \right)^k
\end{aligned}$$

de aquí se obtienen los momentos:

$$\begin{aligned}
E[BN(k,p)] &= G'_{BN(k,p)}(1) = \frac{k}{p} \\
Var[BN(k,p)] &= G''_{BN(k,p)}(1) + G'(1) - (G'^2 = \ldots = \frac{kq}{p^2}
\end{aligned}$$

Proposición El proceso $\{T_k\}_{k\geq 1}$ verifica que:

$$(a) \quad P\{T_{k+1} = n/T_0 \ldots T_k\} = P\{T_{k+1} = n/T_k\}$$

esto es independencia de $T_0, \ldots, T_{k-1}$.

$$(b) \quad P\{T_{k+1} - T_k = m/T_0 \ldots T_k\} = P\{T_{k+1} - T_k = m\} = pq^{m-1}$$

Distribución geométrica.

Ejercicio. Calcular

$$\begin{aligned}
E[T_{k+1} - T_k] &= \frac{1}{p} \\
Var[T_{k+1} - T_k] &= \frac{q}{p^2} \\
G_{T_{k+1}-T_k}(s) &= \frac{ps}{1-q.s}
\end{aligned}$$

Ejercicios de aplicación: Reposición de dispositivos cuya vida $\sim$ geom (p).

$$\pi(m) = pq^{m-1}, \qquad m \geq 1$$

Calcular $P\{T_1 = 3, T_5 = 9, T_7 = 17\}$

Solución:

$$\begin{aligned}
P\{T_1 = 3, T_5 &= 9, T_7 = 17\} = P\{T_1 = 3, T_5 - T_1 = 6, T_7 - T_5 = 8\} \\
&= (\text{independencia } = P(T_1 = 3) P(T_5 - T_1 = 6) P(T_7 - T_5 = 8) \\
& \quad \leftarrow P(T_4 = 6) \rightarrow \quad \leftarrow P(T_2 = 8) \rightarrow \\
&= \begin{pmatrix} 3-1 \\ 1-1 \end{pmatrix} p^1 q^2 \begin{pmatrix} 6-1 \\ 4-1 \end{pmatrix} p^4 q^2 \begin{pmatrix} 8-1 \\ 2-1 \end{pmatrix} p^2 q^6 = 70 p^7 q^{10}
\end{aligned}$$

Calcular $E[T_5/T_3]$, y $E[T_5/T_1 = 3, T_2 = 12, T_3 = 14]$

$$E[T_5/T_3] = E[T_3 + T_5 - T_3/T_3] = T_3 + E[T_2] = T_3 + \frac{2}{p}$$

entonces

$$\begin{aligned}
E[T_5/T_1 = 3, T_2 = 12, T_3 = 14] &= E[T_5/T_3 = 14] \\
&= T_3 + \frac{2}{p} = 14 + \frac{2}{p}
\end{aligned}$$

La vida de cierta componente es una variable aleatoria discreta, cuya distribución es $\pi(m) = pq^{m-1}$, con $m \geq 1$. Al producirse el fallo, se sustituye por otra idéntida. Los reemplazamientos ocurren en los instantes dados por $\{T_k\}_{k\in\mathbb{N}}$. El coste de cada componente nueva es c euros. Se admite la "tasa de descuento" constante para el dinero, esto es a lo largo de n períodos, c euros del instante n tienen un valor $c.\alpha^n$ euros de hoy (obsérvese que $\alpha < 1$ y si la tasa de inflación es de 6 % por ejemplo $\alpha = \frac{1}{1+0{,}06}$). Entonces si cada resposición cuesta c euros, calcúlese la esperanza total del valor de descuento de todas las futuras reposiciones de esta componente.

$$C_k^{(\omega)} = c.\alpha^{Tk(\omega)}, \qquad k \geq 1$$

entonces:

$$C(\omega) = \sum_{k=1}^{\infty} c\alpha^{T_k(\omega)}$$

$$\begin{aligned} E[C] &= \sum_{k=1}^{\infty} cE[\alpha^{T_k}] = \sum_{k=1}^{\infty} c\left(\frac{\alpha\ p}{1-\alpha\ \mathrm{q}}\right)^k \\ &= c\frac{\alpha p}{1-\alpha q}\left(\frac{1}{1-\frac{\alpha p}{1-\alpha q}}\right) = c\frac{\alpha p}{1-\alpha} \end{aligned}$$

ya que :

$$E[\alpha^{T_k}] = [\alpha^{T_1}\alpha^{T_2-T_1}.....\alpha^{T_k-T_{k-1}}] = T_1, T_2-T_1, ...T_k-T_{k-1}$$

tienen la misma distribución geométrica)

$$= G_{geom}(\alpha) = \frac{\alpha p}{1-\alpha q}$$

3. Generalización del Proceso de Sumas Bernoulli

PASEOS ALEATORIOS:

$$X_i = \begin{cases} 1 & p \\ -1 & 1-p=q \end{cases}$$

$$S_n = S_0 + \sum_{i=1}^{n} X_i, \qquad n \in \mathbb{N}$$

Propiedades:

1. Homogeneidad espacial

$$Prob(S_n = j/S_0 = a) = Prob(S_n = j+b/S_0 = a+b)$$

Demostración

Los dos miembros $\equiv \quad Prob(\sum_1^n X_i = j-a)$ $\qquad\square$

2. Homogeneidad temporal

$$Prob(S_n = j/S_0 = a) = Prob(S_{n+m} = j/S_m = a)$$

Demostración

Miembro izquierdo =

$$= P\left(\sum_{i=1}^{n} X_i = j - a\right) = p\left(\sum_{m+1}^{m+n} X_i = j - a\right) =$$

= Miembro derecho.

3. Posee la propiedad de Markov

$$P(S_{n+m} = j/S_0 \ldots S_n) = p(S_{n+m} = j/S_n) \qquad \forall m \geq 0$$

4. Paseo Aleatorio Simétrico (caso $p = 1 = 1/2$)

$$E[S_n] = 0 \qquad \forall n$$

$$Var[S_n] = n \qquad \forall n$$

$$G_{Sn}(s) = (sp + q)^n \quad \forall n$$

Si $X_i \quad i = 1 \ldots n$, es sucesión aleatoria con múltiples valores en $\mathbb{Z}$ para X_i, el proceso suma S_n origina una Cadena de Markov de parámetro discreto.