Ejercicio de Probabilidad

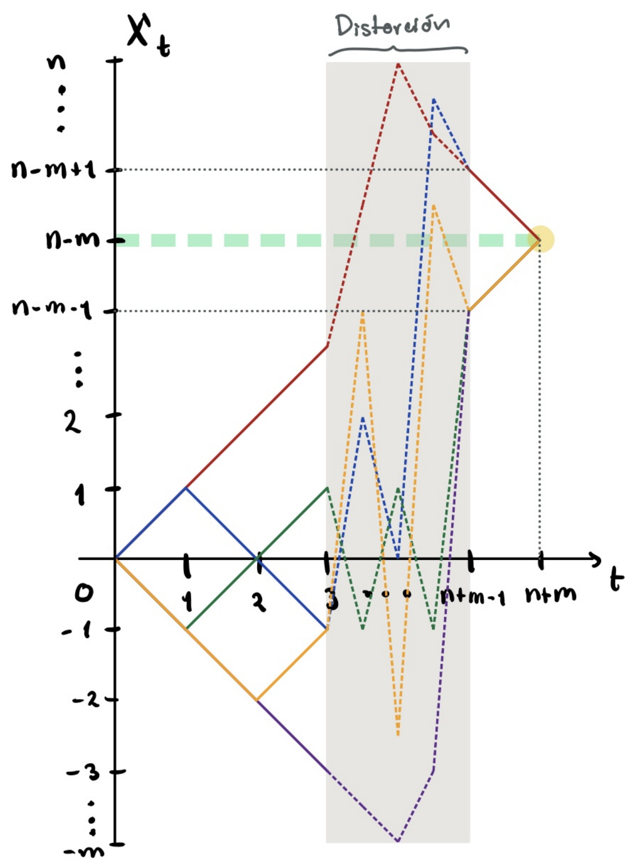
Edgar Alarcón

10/10/2020

El candidato A obtiene n votos y B obtiene m con $n > m$. ¿Cuál es la probabilidad que A siempre vaya en la delantera en los conteos?

Solución.

Sea $\{X_t\}_{t=1}^{n+m}$ la ventaja que tiene el candidato A sobre el candidato B, en este sentido, t denota el número de votante que ejerce sufragio. Pensamos entonces que $t \geq 0$ y definimos $X_0 \equiv 0$. Hagamos una observación de cómo podrían ser las trayectorias.



El área gris denota una posible trayectoria que hay desde $t > 3$ hasta $t < n + m - 1$. Notemos que este es un proceso estocástico pensado a tiempo discreto. Hagamos un par de observaciones:

Observación 1: Toda trayectoria comienza en cero y termina en $n - m > 0$ esto pues tenemos la hipótesis $n > m$.

Observación 2: El punto más alto que puede tener la trayectoria es n, mientras que el más bajo es $-m$. Esto claro, en caso de que los primeros votantes ejerzan sufragio para el mismo votante consecutivamente.

Derivado de lo anterior, tenemos dos principales consecuencias.

Consecuencia 1: Toda ruta en la que exista t_0 tal que $X_{t_0} < 0$ implica la existencia de un t_1 tal que $X_{t_1} \equiv 0$. Esto pues si en algún momento va perdiendo el candidato A, debe haber un momento donde haya un empate ya que sabemos que al final él remontará.

Consecuencia 2: Toda ruta que comience con un votante para B, debe tener un t_1 tal que $X_{t_1} \equiv 0$. Esto de nuevo, por la consecuencia anterior.

Todo esto nos ayuda a comprender mejor cómo se está comportando la ruta de nuestro proceso. No debemos perder de vista el objetivo del ejercicio, se pueden hacer observaciones igualmente importante como que si $t > 2m \Rightarrow X_t > 0$, pero con lo anterior podemos hacer frente a lo siguiente.

Afirmación 3: "$\exists t$ tal que $X_t \equiv 0$" $\Leftrightarrow$ "$\exists t$ tal que $X_t \leq 0$"

Demostración.

La primera implicación ($\Rightarrow$) es clara, ya que $\leq$ es en particular $=$. La segunda ($\Leftarrow$) se da gracias a lo establecido en la *Consecuencia 1*.

$\square$

Afirmación 4: El evento "$X_t > 0\ \forall t$" (que significa que el candidato A vaya siempre a la delantera) tiene como evento complementario "$\exists t$ tal que $X_t \equiv 0$" (que significa que en algún momento exista un empate).

Demostración.

El evento complementario a "$X_t > 0\ \forall t$" es directamente que "$\exists t$ tal que $X_t \leq 0$" el cual es equivalente a "$\exists t$ tal que $X_t \equiv 0$" gracias a lo establecido en la *afirmación 3*.

$\square$

Afirmación 5: "$X_1 = -1$ y $\exists t$ tal que $X_t \equiv 0$" $\Leftrightarrow$ "$X_1 = -1$".

Demostración.

Gracias a lo establecido en la *consecuencia 2* tenemos que "$X_1 = -1$" $\Rightarrow$ "$\exists t$ tal que $X_t \equiv 0$". De tal manera que pensando los eventos como conjuntos $\{X_1 = -1\} \subset \{\exists t \text{ tal que } X_t \equiv 0\}$. De tal manera que $\{X_1 = -1\} \cap \{\exists t \text{ tal que } X_t \equiv 0\} = \{X_1 = -1\}$

$\square$

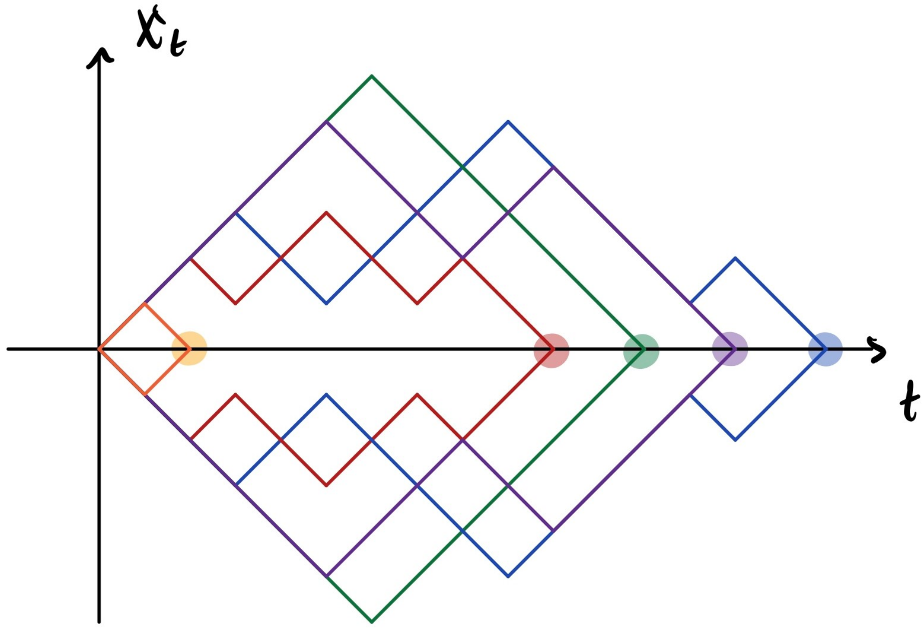
Afirmación 6: $\mathbb{P}[X_1 = 1, \exists t \text{ tal que } X_t \equiv 0] = \mathbb{P}[X_1 = -1]$

Demostración.

Probar esto, debido a la *Afirmación 5* es equivalente a probar:

$$\mathbb{P}[X_1 = 1, \exists t \text{ tal que } X_t \equiv 0] = \mathbb{P}[X_1 = -1, \exists t \text{ tal que } X_t \equiv 0]$$

Esta probabilidad es equivalente pues toda ruta tal que $X_1 = 1$ y toca el cero, tiene su evento simétrico tal que $X_1 = -1$ y toca el cero; esto pues, podemos conmutar las votaciones en un orden adecuado para que la probabilidad se mantenga. En otras palabras, no hay camino que comience en 1 y toque el cero, que no pueda tener un fragmento simétrico dado por uno que empiece en -1. Un ejemplo visual de esto lo podemos notar en el siguiente gráfico



Nota: Posterior al primer punto en que toca el cero, ambos caminos pueden ser completamente idénticos pues han votado la misma cantidad de personas.

□

Habiendo hecho todo lo anterior, podemos comenzar con la prueba del ejercicio principal.

$$\begin{aligned}
\mathbb{P}[X_t > 0 \quad \forall t \geq 1] &= 1 - \mathbb{P}[\exists t \quad \text{tal que} \quad X_t \leq 0] \\
&= 1 - (\mathbb{P}[\exists t \quad \text{tal que} \quad X_t \leq 0, X_1 = -1] + \mathbb{P}[\exists t \quad \text{tal que} \quad X_t \leq 0, X_1 = 1]) \\
&= 1 - \mathbb{P}[X_1 = -1] \qquad \text{(Af 3, 4 y 5)} \\
&= 1 - 2\left(\frac{m}{n+m}\right) = \frac{n-m}{n+m}
\end{aligned}$$

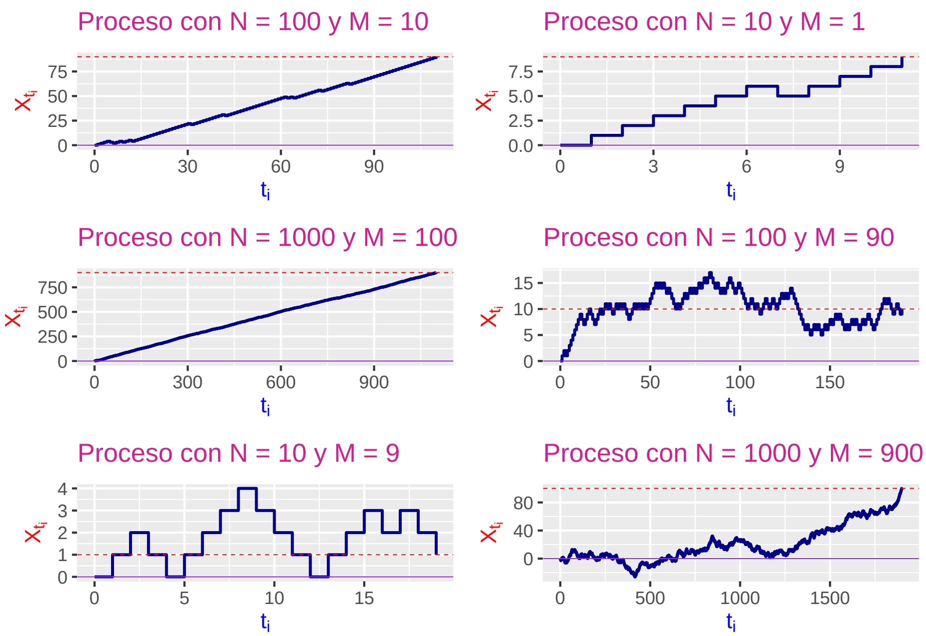
$$\therefore \mathbb{P}[X_t > 0 \quad \forall t \geq 1] = \frac{n-m}{n+m}$$

A continuación mostramos un código que muestra rutas para diferentes valores de n y m. Mostraremos únicamente el algoritmo de simulación.

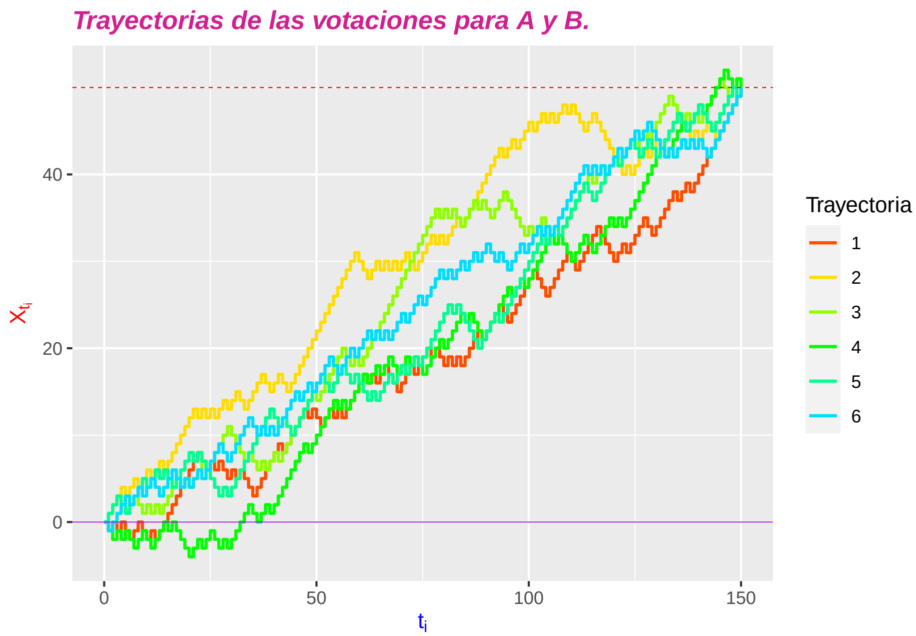
```
Sim_Trayectorias<-function(argumentos){


  # Votantes para A
  N <- argumentos[1]
  # Votantes para B
  M <- argumentos[2]
  # Número de tiempos (contando el cero)
  n = N+M+1 # EN ESTE CASO, el número de tiempos va ligado con los parámetros N y M.

  # Los tiempos van entre 0:(N+M) pues son todos los votantes y el inicio en cero.
  times <- 0:(N+M)
```

```r
  # Algoritmo de simulación
  algoritmo<-function(pasado,t){
    # Veamos el voto.
    voto<-sample(x = c(1,-1),  # Si vota A aumenta, B disminuye.
                 size = 1,      # Emite un voto
                 prob = c(N,M)/(N+M)) # Probabilidad de votar A y B respectivamente.
    # Cuando los votantes disminuyen conforme van pasando.
    # (Obligamos a que la N y M sean guardadas aún fuera de la función)
    if(voto==1){
      N<<-N-1 # Disminuyen los votantes para A
    }else{
      M<<-M-1 # Disminuyen los votantes para A
    }
    # Movamos la trayectoria
    presente <- pasado + voto
    return(presente)
  }


  # Simulamos las trayectorias
  xt1<- 0 # Xt1 = X0 = 0
  Xt <- Reduce(f = algoritmo,                     #Aplica este algoritmo.
               x = times[2:n]-times[1:(n-1)], #Usa estos tiempos.
               init = xt1,                    #Aquí partimos.
               accumulate = TRUE)             #Toda la trayectoria.

  return(Xt)

}

# Realizamos las simulaciones:
argumentos<-list(c(100,10),c(10,1),c(1000,100),c(100,90),c(10,9),c(1000,900))
set.seed(21)
Xts <- sapply(X = argumentos,FUN = Sim_Trayectorias)
```

Adicionalmente vemos los gráficos



Podemos también observar rutas que tengan el mismo parámetro simplemente cambiando un poco el código y lo hacemos para $n = 100$ y $m = 50$.



Finalmente con todo esto, podemos hacer muchas rutas y encontrar una probabilidad empírica y compararla contra la propuesta

```
# Si se desea modificar los parámetros hágalo aquí ~~~~~~~~~~~~~~~~~
# Votantes para A
N = 100
# Votantes para B
M = 50

# Número de tiempos (contando el cero)
```

```r
n = N+M+1 # EN ESTE CASO, el número de tiempos va ligado con los parámetros N y M.

# Los tiempos van entre 0:(N+M) pues son todos los votantes y el inicio en cero.
times <- 0:(N+M)

# Realizamos las simulaciones:
num_trayec = 1000
set.seed(6)
Xts <- replicate(n = num_trayec,expr = Sim_Trayectorias())
Xts<-Xts[-1,]
# Función que revisa si alguno es negativo por
any_neg<-function(x){
  ifelse(any(x<=0),"No siempre positivos","Siempre positivos")
}
ensayos <- apply(X = Xts,MARGIN = 2,any_neg)
```

y de esta manera estimamos la probabilidad

```r
# Probabilidades Empíricas
table(ensayos)/length(ensayos)
```

```
## ensayos
## No siempre positivos    Siempre positivos
##                0.661                0.339
```

```r
# Probabilidad Teórica.
(N-M)/(N+M)
```

```
## [1] 0.3333333
```

Donde comprobamos que en efecto se satisface nuestra proposición.