```r
# POLYA RANDOM WALK
n <- 1000 # number of iterations
alpha0 <- 1   # the parameter alpha0 of the Polya process
alpha1 <- 10 # the parameter alpha1 of the Polya process
alpha <- alpha0+alpha1 # the parameter alpha of the Polya process
m0 <- 0  # the initial number of failures
m1 <- 0 # the initial number of successes
Y <- rep(0,times=n)# definition of the r.v. Y
# start the main cycle
for (m in 0:(n-1)){
  if ( runif(1) < (alpha1+m1)/(alpha+m)) { # comparison with the predictive probability of success
    Y[m+1] <- +1 # upward jump
    m1 <- m1+1 # update total number of successes
  } # end if
  else { # else a failure
    Y[m+1] <- -1 # downward jump
    m0 <- m0+1 # update total number of failures
    } # end else
} # end for
S <- cumsum(Y) # implementing Eq.(46) of notes Scalas2 (definition of Polya random walk)
t <- c(1:n) # time steps
plot(t,S,type="l",xlab="m",ylab="S(m)",main="Polya random walk")
```

```r
##########################################
# PROGRAM POLYA COMPARISON WITH BINOMIAL #
##########################################

##############
# Parameters #
##############
N<-10000 #number of realizations
m<-10 #total number of individuals/observations
alpha0<-100 #parameter of category 0 (failure)
alpha1<-100 #parameter of category 1 (success)
alpha<-alpha0+alpha1 #parameter alpha
#
M1<-rep(c(0),times=(m+1)) #vector storing MC results
#
for(i in 1:N){ #cycle of realizations
  m0<-0 #initialization of the number of failures
  m1<-0 #initialization of the number of successes
  #
  for(k in 1:m){#cycle on individuals/observations
    if ( runif(1) < (alpha1+m1)/(alpha+k) ) { #predictive probability for success
      m1<-m1+1 #update of successes number
    }#end if
    else {
      m0<-m0+1 #update of failures number
    }#end else
  }#end for (k)
  #
  M1[m1+1]<-M1[m1+1]+1 #storing the number of successes
```

```
  #M1[k+1] at the end of the for cycle gives the number of iterations in which k successes happened
}#end for (i)
M1[1:(m+1)] #displays simulation values
k<-seq(0,10) #creates vector of integers from 0 to 10
N*dbinom(k,m,0.5) # displays binomial prediction


#######################
# Plots of the values #
#######################
Propor<-(N)*dbinom(k,m,0.5) #proportion of the binomial distribution having parameters m and 0.5
plot(k,M1[1:11], ylab="number of successes", xlab="number of observations") #plot of the number of successes
points(k,Propor,type="p",pch=3) #plot of the mean of the binomial distribution
legend(cex=1,"topright",c("Polya", "Binomial"), pch=c(1,3))
```

Birthday problem

lunedì 14 giugno 2021 13:25

There are n persons in a room, $n < 365$.

<u>Ob</u>: What is the prob. that at least two of them have a common birthday?

$$\begin{cases} n = \text{number of persons in the room} \\ g = 365 \end{cases}$$

$X_i = x_i \longleftrightarrow$ the i-th person is born on the x_i-th day of the year

For ex. $X_i = 1 \leftrightarrow$ the i-th person is born on Jan 1st

$X_i = 365 \leftrightarrow$ the i-th person is born on Dec. 31st

$\rightarrow (\underbrace{X_1 = x_1}_{365 \text{ possibilities}}, \dots, \underbrace{X_n = x_n}_{365 \text{ possibilities}})$ 365^n possible individual descriptions

$\rightarrow (Y_1 = n_1, \dots, Y_g = n_g)$ $Y_i = n_i \longleftrightarrow$ there are n_i persons born on the i-th day of the year

FREQUENCY VECTOR

$\rightarrow (Z_0 = z_0, \ldots, Z_m = z_m)$ $\quad Z_i = z_i \leftrightarrow$ there are z_i days with i birthdays

PARTITION VECTOR

complementary situation (simpler to analyze)

no common birthday $\longleftrightarrow (Z_0 = 365 - n, Z_1 = n, Z_2 = 0, \ldots, Z_n = 0)$

From eq. (6) of the notes Scalas2

$$(*) \quad \underline{P}(Y_1 = n_1, \ldots, Y_g = n_g) = \frac{n!}{n_1! \ldots n_g!} \cdot \underline{P}(X_1 = x_1, \ldots, X_n = x_n)$$

In a similar way, one can show that

$$(**) \quad \underline{P}(Z_0 = z_0, Z_1 = z_1, \ldots, Z_n = z_n) = \frac{g!}{z_0! \cdot z_1! \cdots z_n!} \cdot \underline{P}(Y_1 = n_1, \ldots, Y_g = n_g)$$

From eqs $(*)$ and $(**)$

$$(1) \quad \underline{P}(Z_0 = z_0, \ldots, Z_n = z_n) = \frac{g!}{z_0! z_1! \cdots z_n!} \cdot \frac{n!}{n_1! \ldots n_g!} \cdot \underbrace{\underline{P}(X_1 = x_1, \ldots, X_n = x_n)}_{\text{let's compute this probability}}$$

We assume that $X_1, \dots, X_n$ are i.i.d. random variables and that all the days are equiprobable, so that $\mathbb{P}(X_i = x_i) = \frac{1}{365}$ for any x_i.

Hence

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) \underset{\substack{\downarrow \\ \text{because of the} \\ \text{independence}}}{=} \prod_{i=1}^{n} \mathbb{P}(X_i = x_i) \underset{\substack{\downarrow \\ \text{because of the} \\ \text{identical distribution,} \\ \text{for a fixed } i \in \{1, \dots, n\}}}{=} \left(\mathbb{P}(X_i = x_i)\right)^n = \frac{1}{365^n}.$$

So from **(1)**

$$\mathbb{P}(Z_0 = 365-n,\ Z_1 = n,\ Z_2 = 0, \dots, Z_n = 0) = \frac{365!}{(365-n)!\,\cancel{n!}\,0!\dots 0!} \cdot \frac{\cancel{n!}}{\underbrace{1!\dots 1!}_{n \text{ times}}\ \underbrace{0!\dots 0!}_{\substack{365-n \\ \text{times}}}} \cdot \frac{1}{365^n}$$

$$= \frac{365!}{(365-n)! \cdot 365^n}.$$

Hence, the probability of having at least two common birthdays is

$$p = 1 - \underline{P}\left(Z_0 = 365 - n,\ Z_1 = n,\ Z_2 = 0, \dots, Z_n = 0 \right) =$$

$$= 1 - \frac{365!}{(365-n)!\,365^n}.$$

If $n \geq 23$, then $p > 0.5$ (See the screen below ...)