LECTURE 21: The Bernoulli process

- Definition of Bernoulli process
- Stochastic processes
- Basic properties (memorylessness)
- The time of the kth success/arrival
- Distribution of interarrival times
- Merging and splitting
- Poisson approximation

The Bernoulli process

- A sequence of independent Bernoulli trials, X_i
- At each trial, i:

$$\begin{aligned} \mathbf{P}(X_i = 1) &= \mathbf{P}(\text{success at the } i\text{th trial}) = p \\ \mathbf{P}(X_i = 0) &= \mathbf{P}(\text{failure at the } i\text{th trial}) = 1 - p \end{aligned}$$

$0 < p < 1$

- Key assumptions:
  - Independence
  - Time-homogeneity
- Model of:
  - Sequence of lottery wins/losses
  - Arrivals (each second) to a bank
  - Arrivals (at each time slot) to server
  - ...

- Jacob Bernoulli (1655–1705)

Stochastic processes

infinite

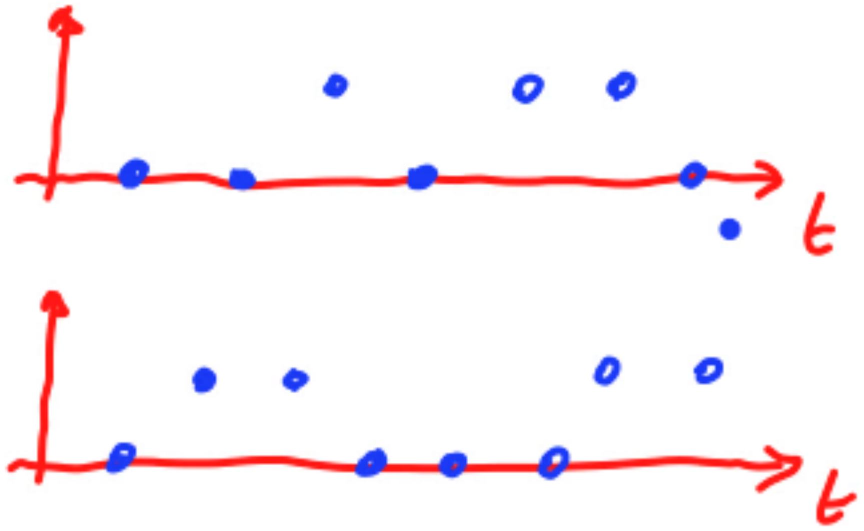
- First view: sequence of random variables $X_1, X_2, \ldots$

Interested in: $\mathbf{E}[X_i] = p$ $\quad$ $\text{var}(X_i) = p(1-p)$ $\quad$ $p_{X_i}(x) = \begin{cases} p & x=1 \\ 1-p & x=0 \end{cases}$

$p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = p_{X_1}(x_1)\cdots p_{X_n}(x_n)$

for all n

- Second view – sample space:



$\Omega =$ set of infinite sequences of 0's and 1's

- Example (for Bernoulli process):

$\mathbf{P}(X_i = 1 \text{ for all } i) = 0$ $\quad$ $(p<1)$

$\leq \mathbf{P}(X_1 = 1, \ldots, X_n = 1) = p^n, \text{ for all } n$

Number of successes/arrivals S in n time slots

- $S = X_1 + \cdots + X_n$
- $\mathbf{P}(S = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, \ldots, n$
- $\mathbf{E}[S] = np$
- $\mathrm{var}(S) = np(1-p)$

Time until the first success/arrival

- $T_1 =$ $\min\{i : X_i = 1\}$

- $\mathbf{P}(T_1 = k) =$ $P(\underbrace{0\,0 \cdots 0}_{k-1}\ 1) = (1-p)^{k-1}\,p$

  $k = 1, 2, \ldots$

- $\mathbf{E}[T_1] = \dfrac{1}{p}$

- $\mathrm{var}(T_1) = \dfrac{1-p}{p^2}$

Independence, memorylessness, and fresh-start properties

$\{X_i\} \sim \text{Ber}(p)$

$Y_1 = X_6$ $\quad X_{n+1}$

$Y_2 = X_7$ $\quad X_{n+2}$

$\vdots$

$\{Y_i\}$ $\quad i = 1, 2, \ldots$

① $\{Y_i\}$ independent of $X_1, \ldots, X_{\cancel{5}\,n}$

② $\text{Ber}(p)$

- Fresh-start after time n

$Y_1 = X_{T_1+1}$

$Y_2 = X_{T_1+2}$

$\vdots$

① $\{Y_i\}$ independent of $X_1, \ldots, X_{T_1}$

② $\text{Ber}(p)$

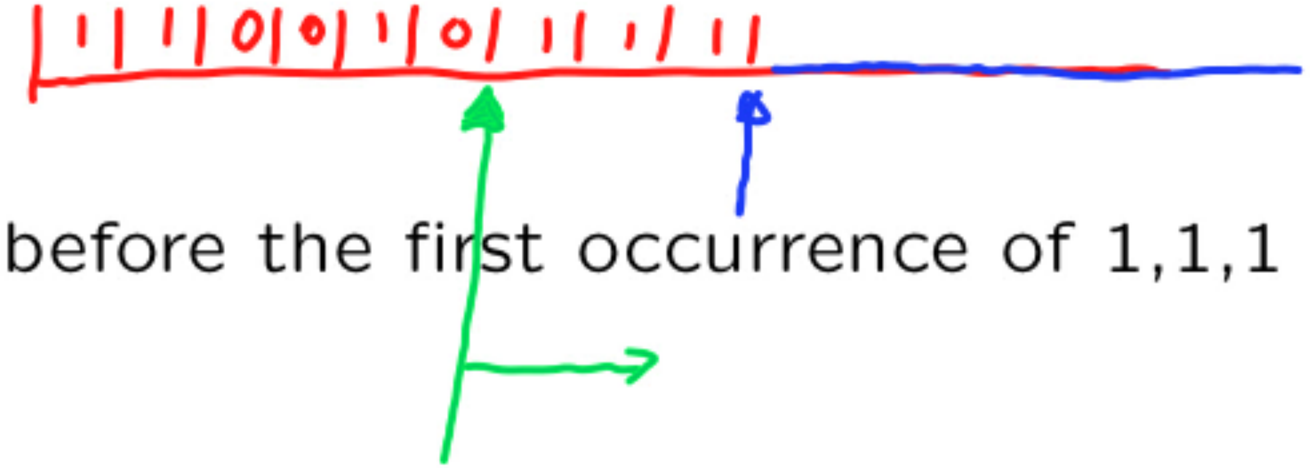
- Fresh-start after time T_1

Independence, memorylessness, and fresh-start properties

- Fresh-start after a random time N ?

  $N =$ time of 3rd success

  $N =$ first time that 3 successes in a row have been observed



  $N =$ the time just before the first occurrence of 1,1,1

N is causally determined

N not causally determined

The process $X_{N+1}, X_{N+2}, \ldots$ is:

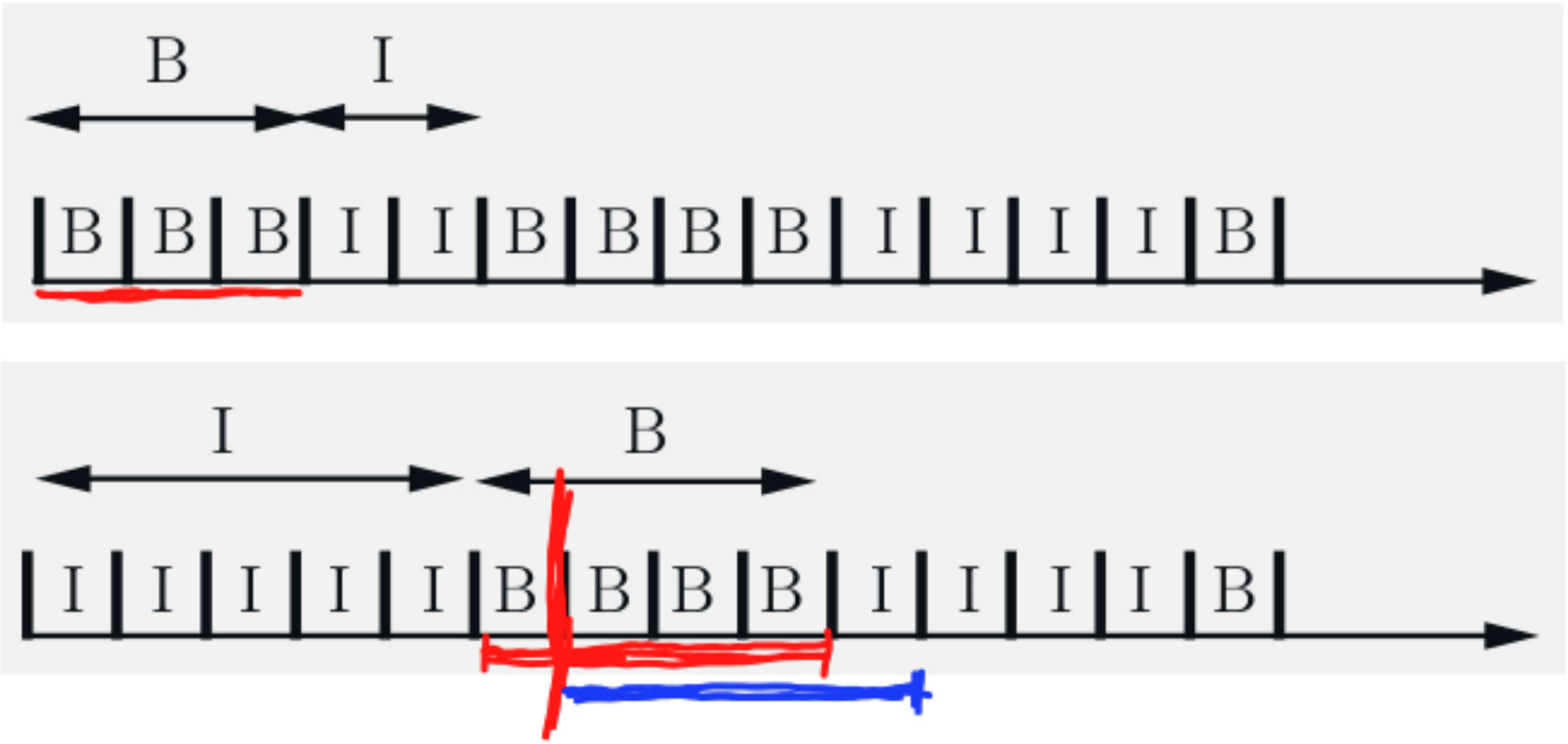
– a Bernoulli process

– independent of $N, X_1, \ldots, X_N$

(as long as N is determined "causally")

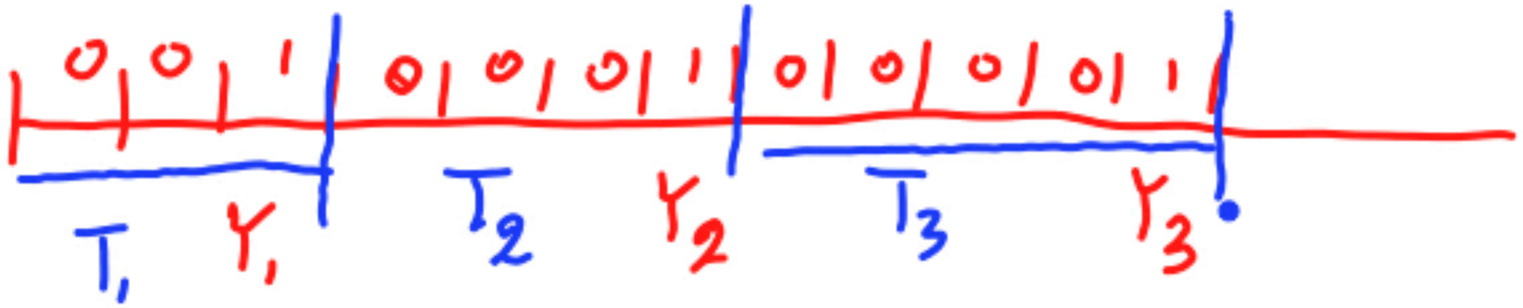
The distribution of busy periods

- At each slot, a server is busy or idle (Bernoulli process) p
- First busy period: $Geo(1-p)$
  - starts with first busy slot
  - ends just before the first subsequent idle slot



$Geo(1-p)$

Time of the kth success/arrival



- Y_k = time of kth arrival

$$Y_k = T_1 + \cdots + T_k$$

- T_k = kth inter-arrival time $= Y_k - Y_{k-1}$ $\quad (k \geq 2)$

- The process starts fresh after time T_1

- T_2 is independent of T_1; Geometric(p); etc.

Time of the kth success/arrival

$$\mathbf{P}(Y_k = t)$$

$$= \mathbf{P}(k-1 \text{ arrivals in time } t-1)$$

$$\cdot\, \mathbf{P}(\text{arrival at time } t)$$

$$= \binom{t-1}{k-1} p^{k-1} (1-p)^{t-k} \cdot p$$

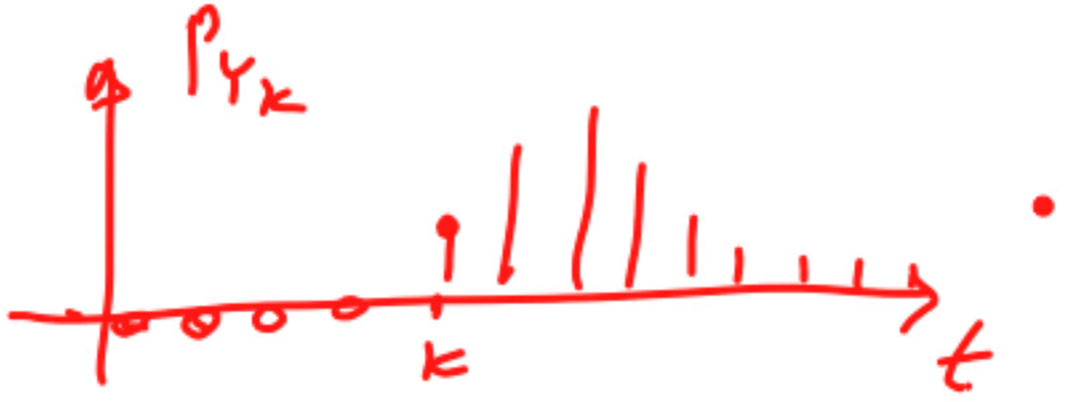
$Y_k = T_1 + \cdots + T_k$

the T_i are i.i.d., Geometric(p)

$\mathbf{E}[Y_k] = \dfrac{k}{p}$ $\quad$ $\text{var}(Y_k) = \dfrac{k(1-p)}{p^2}$

$p_{Y_k}(t) = \dbinom{t-1}{k-1} p^k (1-p)^{t-k},$

$t = k, k+1, \ldots$



Merging of independent Bernoulli processes

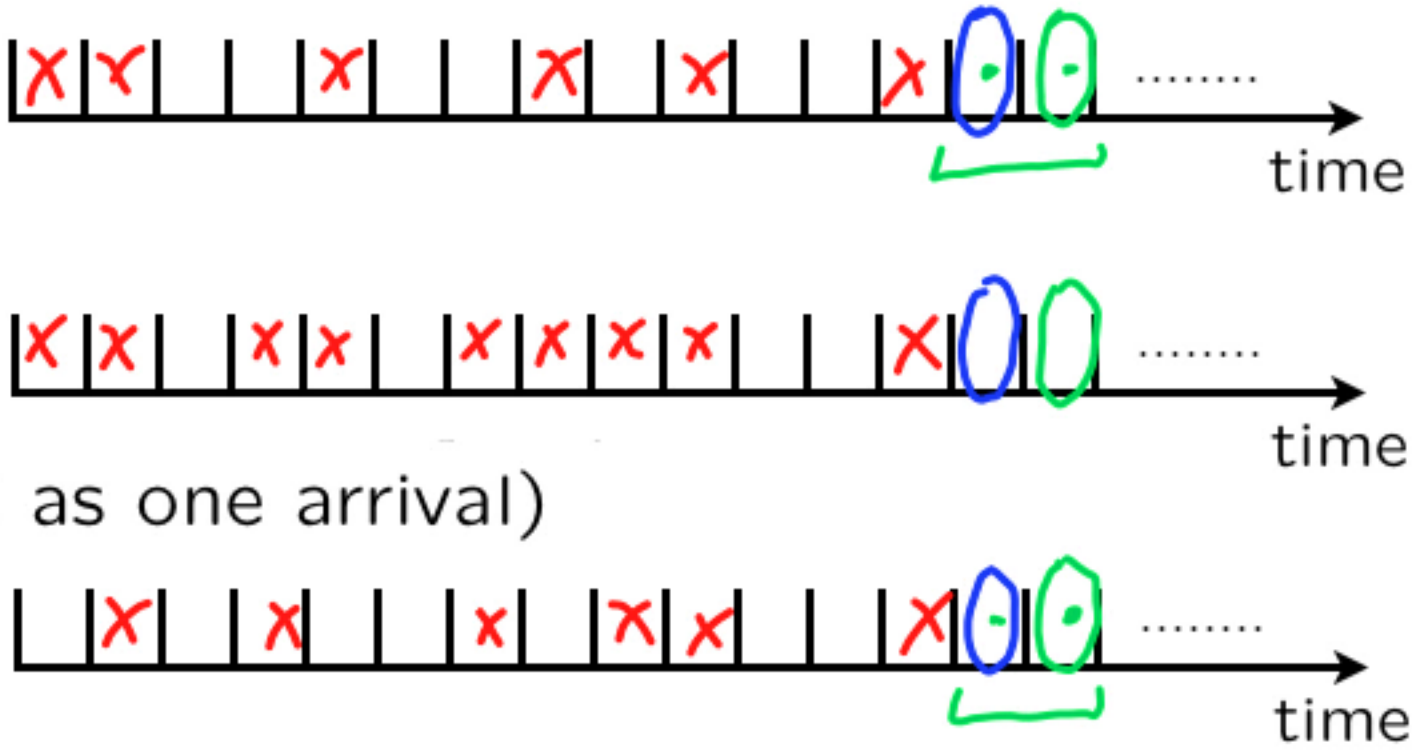
X_t Bernoulli(p)

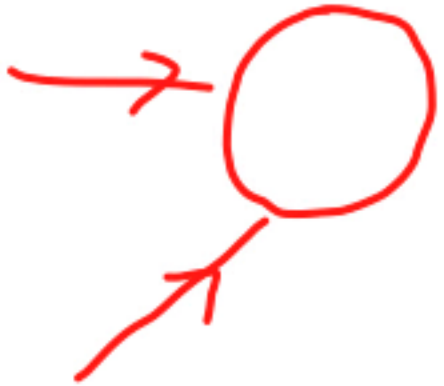
Z_t merged process

Bernoulli($p + q - pq$)

(collisions are counted as one arrival)

Y_t Bernoulli(q)





$Z_t = g(\underline{X_t, Y_t})$ $\quad (Z_1, \ldots, Z_t)$

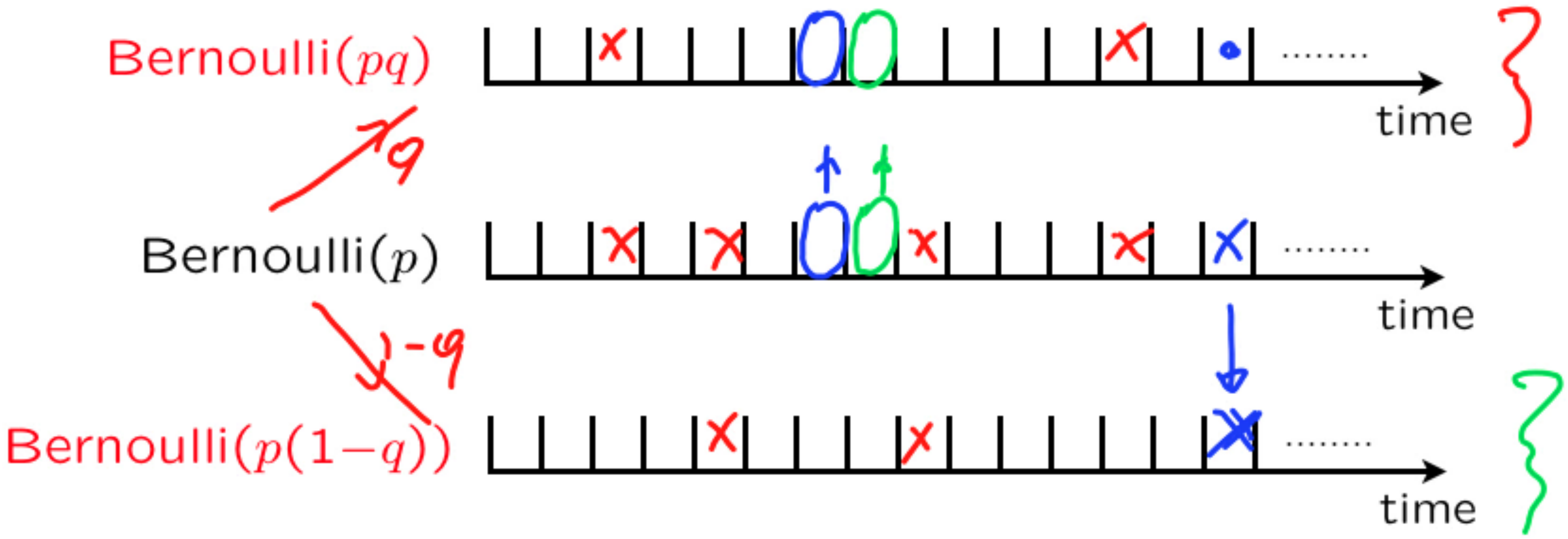
$Z_{t+1} = g(\underline{X_{t+1}, Y_{t+1}})$ $\quad 1 - (1-p)(1-q)$

$\mathbf{P}(\text{arrival in first process} \mid \text{arrival}) = \dfrac{p}{p+q-pq}$

Y \ X	0	1
1	$(1-p)q$	pq
0	$(1-p)(1-q)$	$p(1-q)$

Splitting of a Bernoulli process

- Split successes into two streams, using independent flips of a coin with bias q
  - assume that coin flips are independent from the original Bernoulli process



- Are the two resulting streams independent? No

Poisson approximation to binomial

- Interesting regime: large n , small p, moderate $\underline{\lambda = np}$

$\cdot\ n \to \infty$

$p \to 0 \qquad p = \frac{\lambda}{n}$

- Number of arrivals S in n slots: $\underline{p_S(k)} = \frac{n!}{(n-k)!\,k!} \cdot p^k (1-p)^{n-k}, \quad k = 0, \ldots, n$

> For fixed $k = 0, 1, \ldots,$
>
> $p_S(k) \to \frac{\lambda^k}{k!} e^{-\lambda},$

$$= \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \cdot \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \cdot \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^{n} \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$\xrightarrow[n \to \infty]{} 1 \cdot 1 \cdots 1 \cdot \underbrace{\frac{\lambda^k}{k!}\, e^{-\lambda} \cdot 1}$$

- Fact: $\lim_{n \to \infty} (1 - \lambda/n)^n = e^{-\lambda}$