Ανάλυση της Απόδοσης Πληροφοριακών Συστημάτων

Στοχαστικές Διαδικασίες

◆

Η Διαδικασία Bernoulli

Γιάννης Γαροφαλάκης

Στοχαστικές διαδικασίες

- *Στοχαστική Διαδικασία* (Σ.Δ.): ορίζεται ως μία οικογένεια Τυχαίων Μεταβλητών (Τ.Μ.), ***X(t)***, όπου οι Τ.Μ. ***X*** έχουν δεικτοδοτηθεί με τη χρονική παράμετρο *t*.
- Παράγοντες ταξινόμησης στοχαστικών διαδικασιών
  - *ο χώρος καταστάσεων* (οι τιμές που παίρνουν οι ΤΜ)
    - πεπερασμένες ή αριθμήσιμες τιμές →Σ.Δ. *διακριτών-καταστάσεων* (*αλυσίδα*). *Ο χώρος καταστάσεων* ↔ {0, 1, 2, …}
    - τιμές από ένα πεπερασμένο ή άπειρο συνεχές διάστημα → Σ.Δ. *συνεχών-καταστάσεων*
  - η *χρονική παράμετρος* (επιτρεπτές χρονικές στιγμές αλλαγής κατάστασης)
    - Σ.Δ. Διακριτού-χρόνου [X_n – Στοχαστική Ακολουθία]
    - Σ.Δ. Συνεχούς χρόνου [X(t)]
  - η *στατιστική σχέση* μεταξύ των Τ.Μ.

Στατιστική σχέση μεταξύ των ΤΜ (1)

- Θέλουμε να προσδιορίσουμε την από κοινού PDF στις ΤΜ

  $\vec{X} = [\, X(t_1), X(t_2), \ldots]$, δηλαδή την:

$$F_{\vec{X}}\left(\vec{x}\,;\vec{t}\right) \equiv P\left[X\left(t_1\right) \leq x_1, \; \ldots \, , \; X\left(t_n\right) \leq x_n\right]$$

για όλα τα $\vec{X} = (x_1, x_2, \ldots, x_n)$, $\vec{t} = (t_1, t_2, \ldots, t_n)$ και n.

Στατιστική σχέση μεταξύ των ΤΜ (2)

1. Στάσιμες ΣΔ

   Αμετάβλητες στις ολισθήσεις στο χρόνο. Δηλαδή για οποιοδήποτε σταθερό *τ*, πρέπει: $F_{\vec{X}}\left(\vec{x};\vec{t}+\tau\right)=F_{\vec{X}}\left(\vec{x};\vec{t}\right)$

   όπου $\vec{t} + \tau = (t_1 + \tau, t_2 + \tau, \ldots, t_n + \tau)$.

2. Ανεξάρτητες ΣΔ

   Οι πιο απλές. Δεν υπάρχει καμία δομή ή εξάρτηση των Τ.Μ. τους:

$$f_{\vec{X}}\left(\vec{x};\vec{t}\right) \equiv f_{X_1\ldots X_n}\left(x_1,\ldots,x_n;t_1,\ldots,t_n\right) = f_{X1}\left(x_1;t_1\right)\ldots f_{Xn}\left(x_n;t_n\right)$$

Στατιστική σχέση μεταξύ των ΤΜ (3)

3. Διαδικασίες Markov

Για μια Αλυσίδα Markov $\{X(t)\}$, η πιθανότητα ότι η επόμενη τιμή $X(t_{n+1})$ θα είναι ίση με $\boldsymbol{x_{n+1}}$, εξαρτάται μόνο από την παρούσα τιμή $X(t_n) = \boldsymbol{x_n}$ και όχι από οποιαδήποτε προηγούμενη (ΙΔΙΟΤΗΤΑ ΑΜΝΗΣΙΑΣ).

Ιδιότητα Markov (για αλυσίδα Markov):

$$P\left[X\left(t_{n+1}\right)=x_{n+1} | X\left(t_n\right)=x_n, X\left(t_{n-1}\right)=x_{n-1}, \ldots, X\left(t_1\right)=x_1\right]=$$

$$=P\left[X\left(t_{n+1}\right)=x_{n+1} | X\left(t_n\right)=x_n\right]$$

όπου $t_1 < t_2 < \ldots < t_n < t_{n+1}$, ενώ τα x_i περιέχονται σε κάποιο διακριτό χώρο καταστάσεων.

- Αποδεικνύεται ότι ο **χρόνος παραμονής** σε μια κατάσταση ακολουθεί την:

  Εκθετική Κατανομή (διαδικασία συνεχούς χρόνου), ή την - ισοδύναμη -

  Γεωμετρική Κατανομή (διαδικασία διακριτού χρόνου).

Στατιστική σχέση μεταξύ των ΤΜ (4)

4. Διαδικασίες Γεννήσεων - Θανάτων

   Κλάση των Διαδικασιών Markov: Οι αλλαγές κατάστασης γίνονται μόνο μεταξύ *γειτονικών* καταστάσεων.

   Δηλαδή αν $X(t_n) = i$, τότε $X(t_{n+1}) = i - 1$ ή $X(t_{n+1}) = i + 1$ μόνο.

5. Διαδικασίες Semi Markov
   - Επιτρέπουμε αυθαίρετη κατανομή του χρόνου που η διαδικασία μπορεί να παραμείνει σε μια κατάσταση.
   - Η διαδικασία συμπεριφέρεται σαν Markov κατά τις χρονικές στιγμές αλλαγής κατάστασης, και στην πραγματικότητα *σε αυτές τις στιγμές* λέμε ότι έχουμε μια *συμπυκνωμένη (embedded) αλυσίδα Markov*.
   - Υπερσύνολο των διαδικασιών Markov.

Στατιστική σχέση μεταξύ των ΤΜ (5)

6. Τυχαίοι περίπατοι

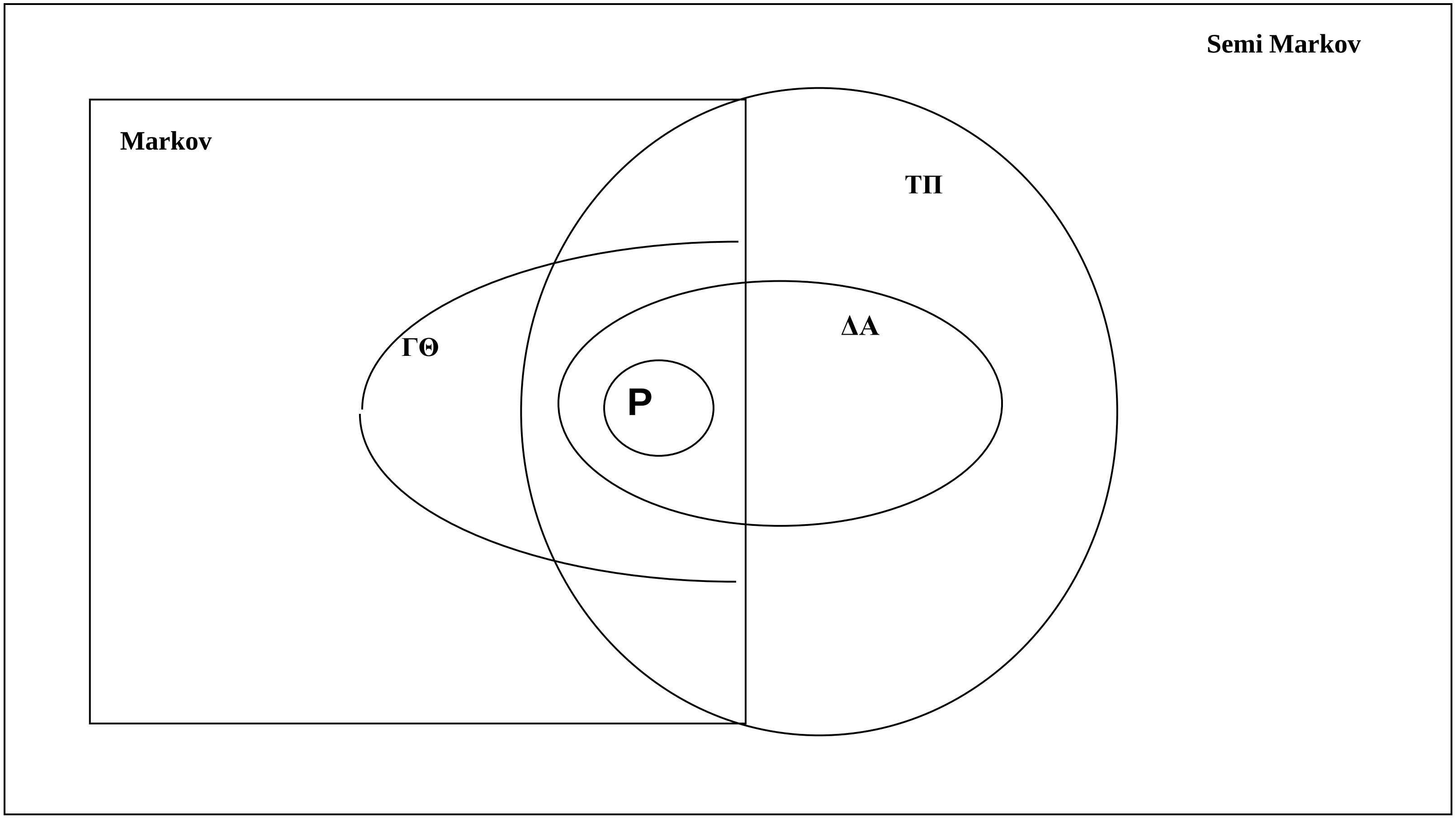
Η επόμενη θέση είναι ίση με την προηγούμενη θέση, συν μια τυχαία μεταβλητή

Δηλαδή, μια ακολουθία ΤΜ *{Sn}* είναι τυχαίος περίπατος αν: $S_n = X_1+X_2+ \ldots +X_n$ όπου $n = 1, 2, \ldots,$ $S_0 = 0$ και $X_1 , X_2 ,\ldots$ είναι ακολουθία ανεξάρτητων ΤΜ με κοινή κατανομή.

7. Διαδικασίες ανανέωσης

- Ειδική περίπτωση των τυχαίων περιπάτων.
- S_n είναι τώρα η ΤΜ που καθορίζει τη *χρονική στιγμή* στην οποία γίνεται η n-οστή μεταβολή κατάστασης και $\{X_n\}$ είναι ένα σύνολο ανεξάρτητων, όμοια κατανεμημένων ΤΜ, όπου η X_n αντιπροσωπεύει το χρόνο μεταξύ της $(n-1)$-οστής και n-οστής μεταβολής κατάστασης . Οι μεταβολές γίνονται μόνο μεταξύ γειτονικών καταστάσεων .

Σχέσεις των κλάσεων Στοχαστικών Διαδικασιών



P: Poisson

Διαδικασία Bernoulli

Μία διαδικασία Bernoulli είναι μία πεπερασμένη ή άπειρη ακολουθία ανεξάρτητων τυχαίων μεταβλητών $X_1, X_2, X_3, \ldots$, τέτοια ώστε

- Για κάθε i η X_i παίρνει την τιμή 0 ή 1.
- Για κάθε i η πιθανότητα το X_i να είναι 1 είναι πάντα ίση με p.

Κατά συνέπεια, η διαδικασία Bernoulli είναι μία ακολουθία ανεξάρτητων ομοιόμορφα κατανεμημένων **δοκιμών Bernoulli**.

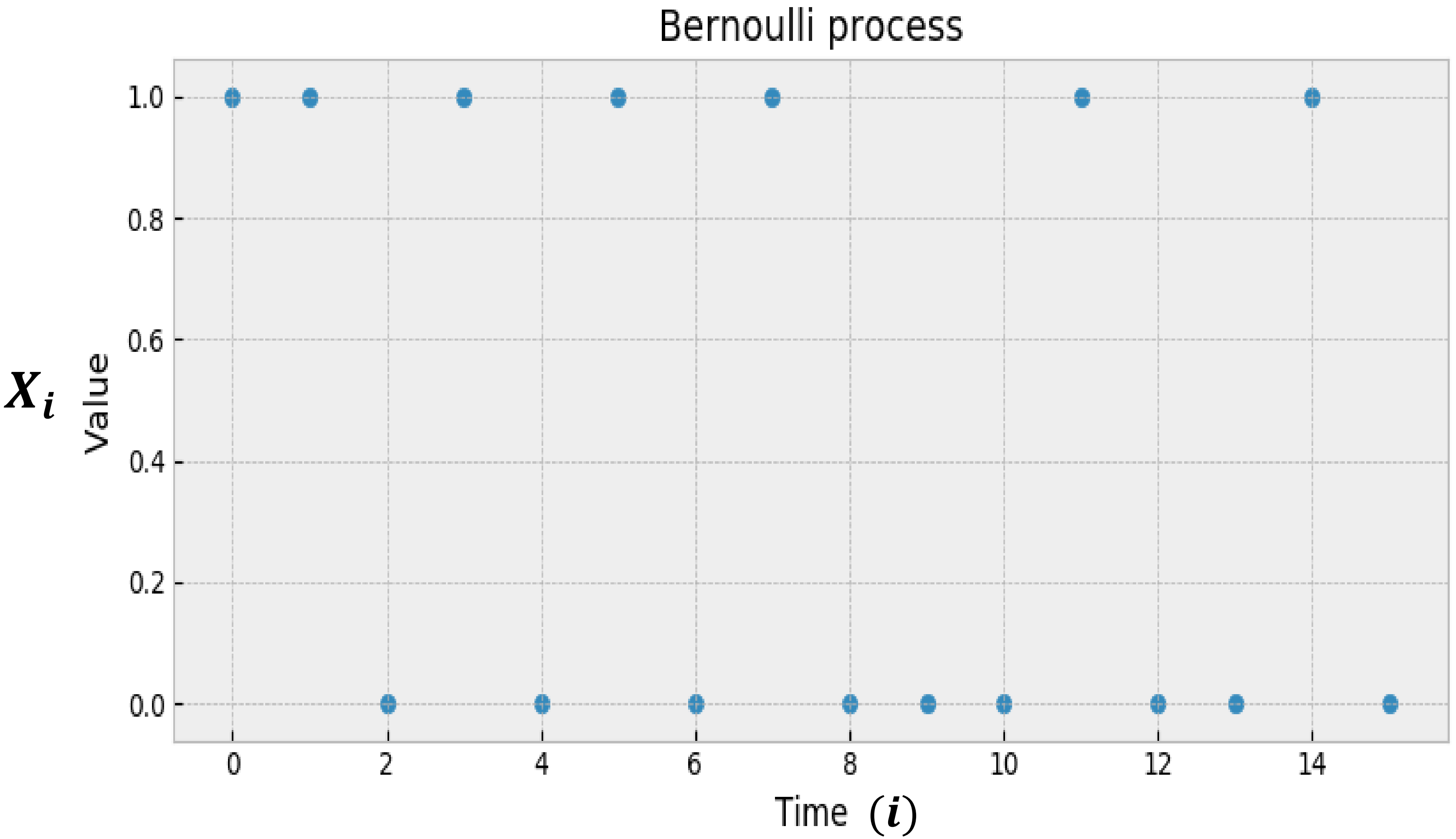
Example

- Μία σειρά ρίψεων ενός νομίσματος
- Αφίξεις πακέτων σε έναν router

Βασικές Ιδιότητες Διαδικασιών Bernoulli

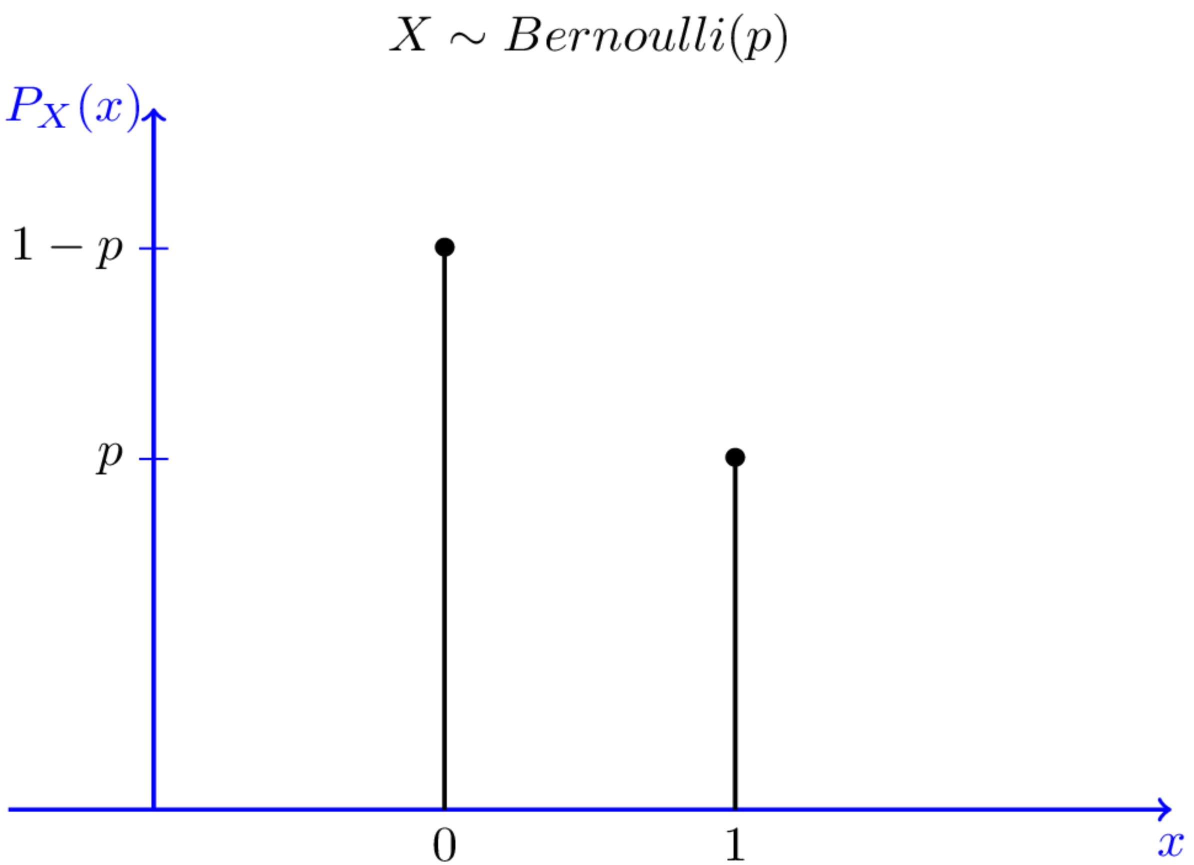
- **Ανεξαρτησία**
- **Έλλειψη μνήμης / Ιδιότητα της Επανεκκίνησης**

Διαδικασία Bernoulli (2)



Κατανομή Bernoulli

- **pmf**: $p_X(x) = \Pr(X = x) = \begin{cases} 1 - p & x = 0 \\ p & x = 1 \end{cases}$

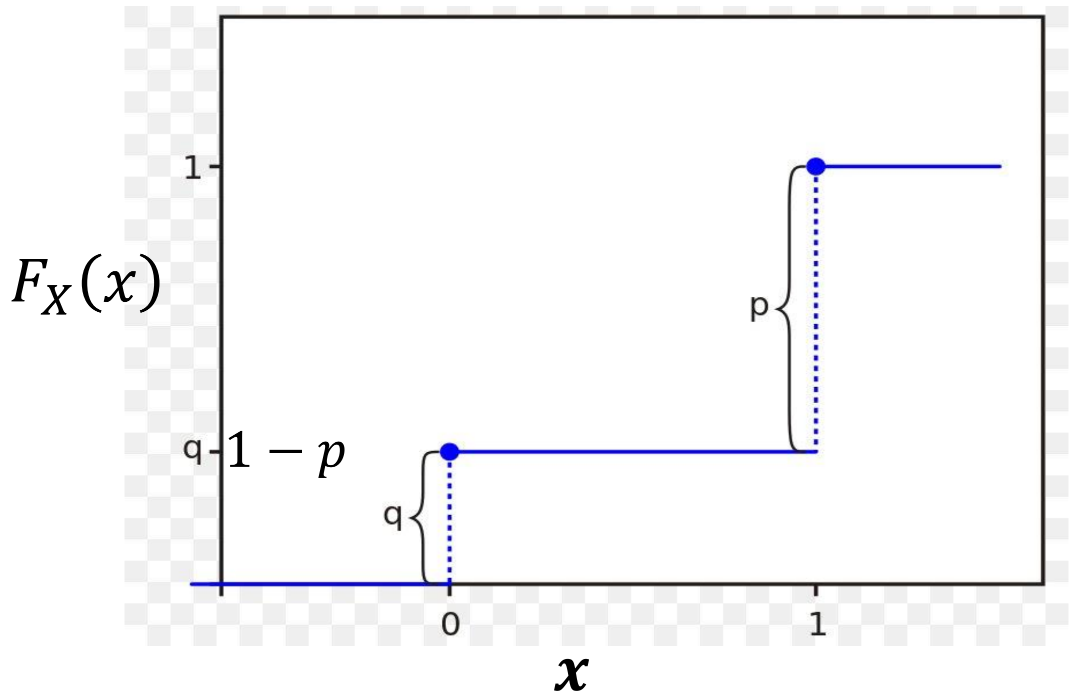


Κατανομή Bernoulli (2)

pmf: $p_X(x) = \Pr(X = x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \end{cases}$

$$F_X(x) = \sum_{x_i \le x} p_X(x_i)$$

- **CDF**: $F_X(x) = \Pr(X \le x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$



Κατανομή Bernoulli (3)

pmf: $p_X(x) = \Pr(X = x) = \begin{cases} 1-p & x = 0 \\ p & x = 1 \end{cases}$

$$\mathrm{E}[X] = \sum_i x_i p_X(x_i)$$

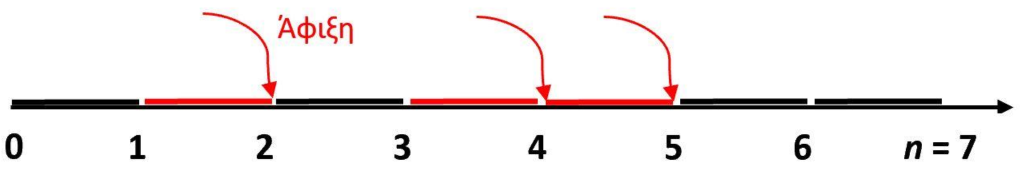
- $E[X] = 0 \cdot (1-p) + 1 \cdot p = \boldsymbol{p}$

$$\mathrm{Var}(X) = \sum_i (x_i - E[X])^2 p_X(x_i)$$

- $Var(X) = (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p = \boldsymbol{p \cdot (1-p)}$

Σχετικές Τυχαίες Μεταβλητές (1)

- Η τ.μ. *N* που μετρά το πλήθος των αφίξεων (επιτυχιών) *k*, κατά τη διάρκεια *n* slots (δοκιμών):



Αριθμός τρόπων κατανομής *k* επιτυχιών σε *n* προσπάθειες:

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$ οπότε:

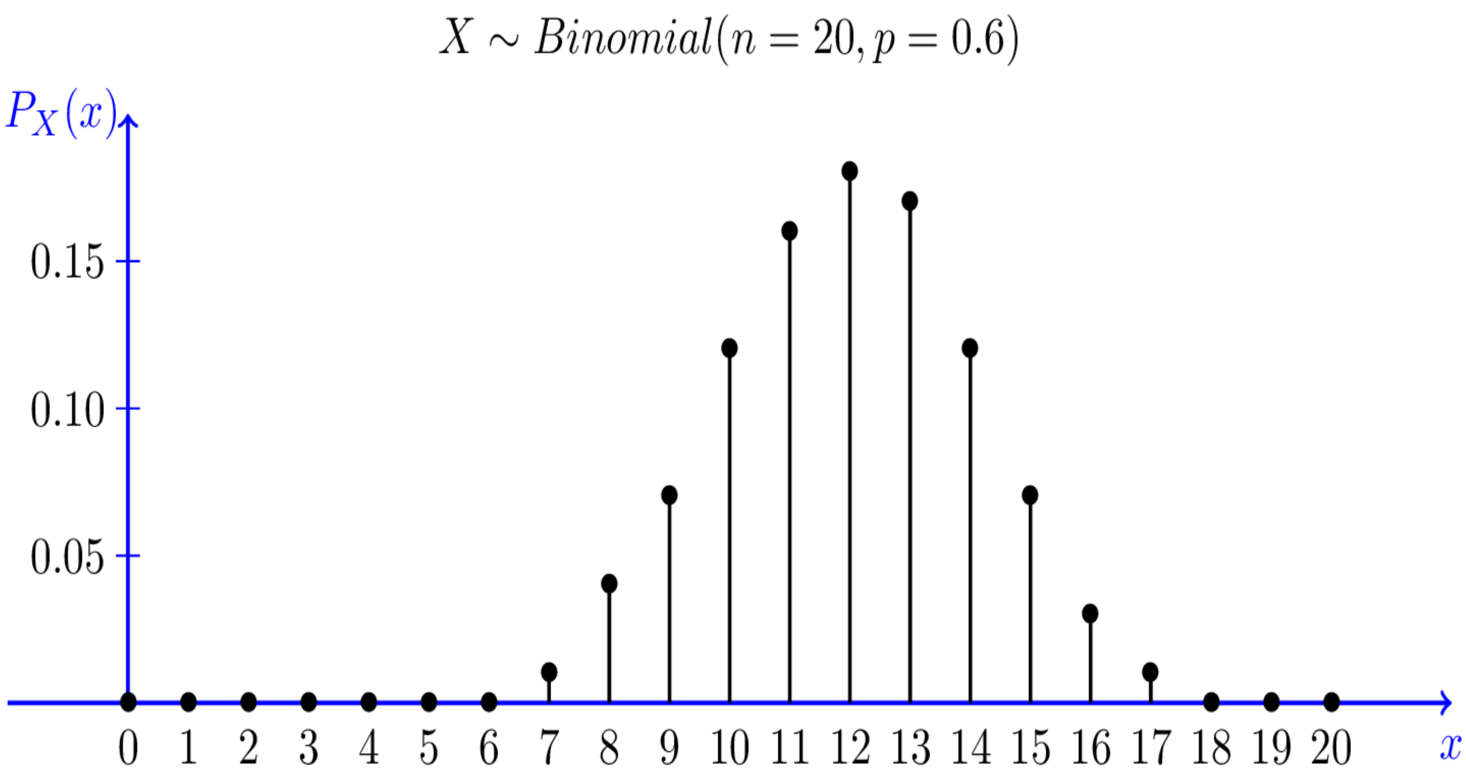
$\Pr\{k \text{ αφίξεις κατά τη διάρκεια } n \text{ slots}\} = p_N(k) =$

$= \boldsymbol{\binom{n}{k} p^k (1-p)^{n-k}}, \quad 0 \le k \le n$ **Binomial** $\mathbf{B}(\boldsymbol{n}, \boldsymbol{p})$

$\mathrm{E[N]} = \boldsymbol{n \cdot p}$ $Var(N) = \boldsymbol{n \cdot p \cdot (1-p)}$

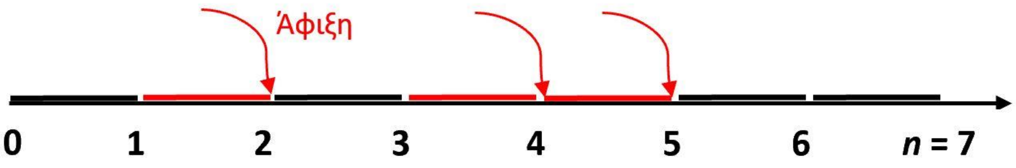
pmf της Binomial B$(\boldsymbol{n}, \boldsymbol{p})$

$$p_N(k) = \binom{n}{k} p^k (1-p)^{n-k}$$



Σχετικές Τυχαίες Μεταβλητές (2)

- Η τ.μ. *T* που μετρά το χρόνο $\boldsymbol{t}$ μεταξύ διαδοχικών αφίξεων:



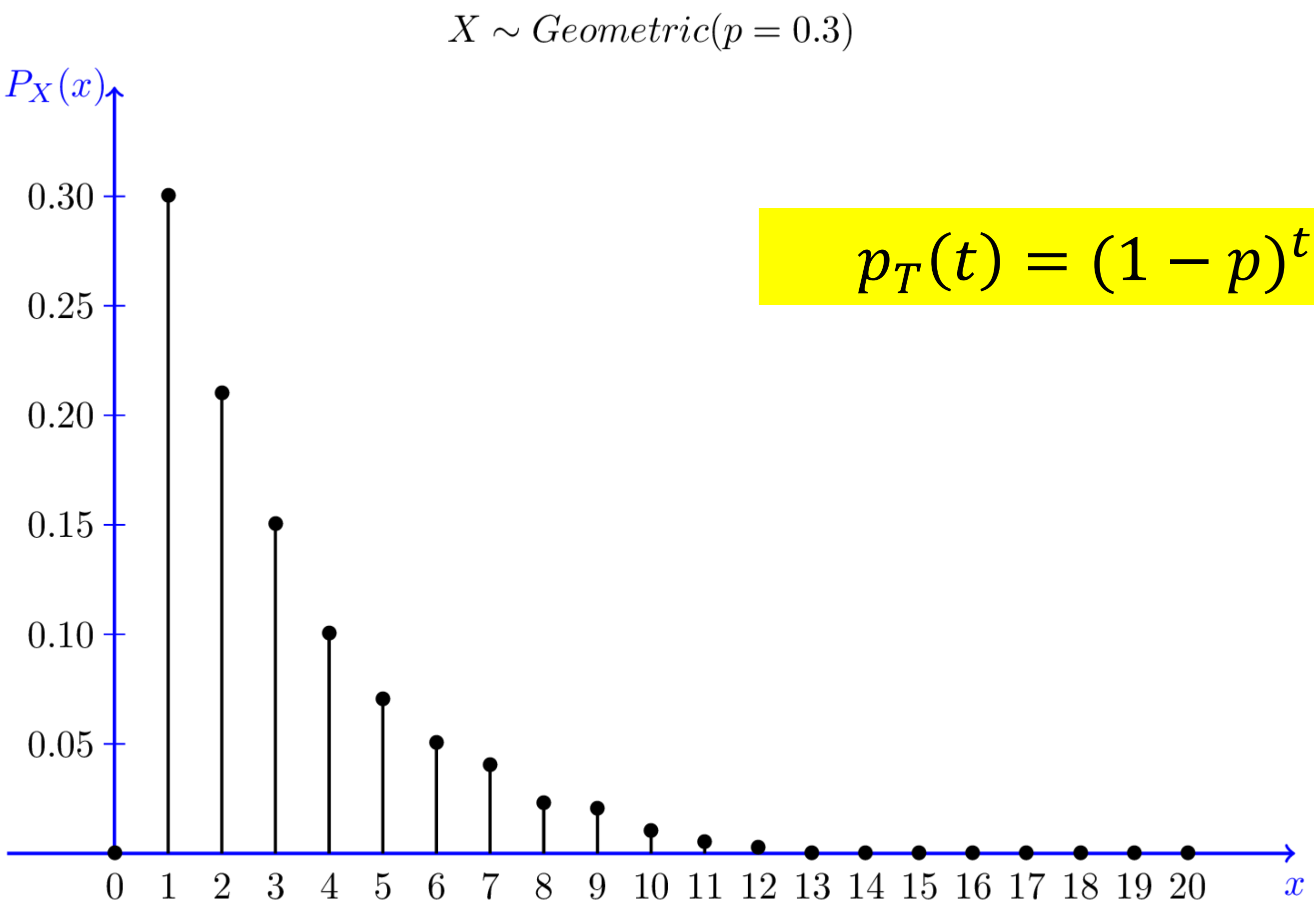
$\Pr\{t$ slots μέχρι την επόμενη άφιξη$\} = p_T(t) =$

$= \boldsymbol{(1-p)^{t-1} \cdot p}$ **Γεωμετρική $\mathbf{NB}(\mathbf{1}, \boldsymbol{p})$**

$\mathrm{E}[\mathrm{T}] = \mathbf{1/p}$

$Var(T) = (\mathbf{1} - \boldsymbol{p})/\boldsymbol{p}^{\mathbf{2}}$

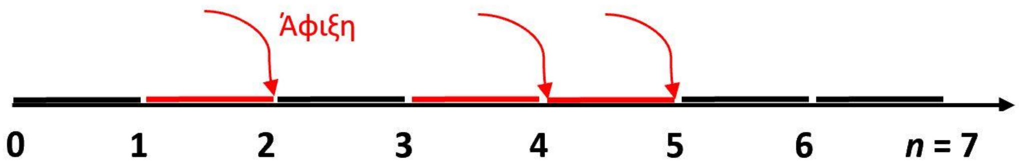
pmf της Γεωμετρικής $\mathbf{NB}(\mathbf{1}, \boldsymbol{p})$



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

Σχετικές Τυχαίες Μεταβλητές (3)

- Η τ.μ. Y_k που μετρά το χρόνο $\boldsymbol{t}$ μέχρι την *k*-οστή άφιξη:



Δηλαδή: $k-1$ αφίξεις σε $t-1$ slots (binomial), και μετά, 1 άφιξη.

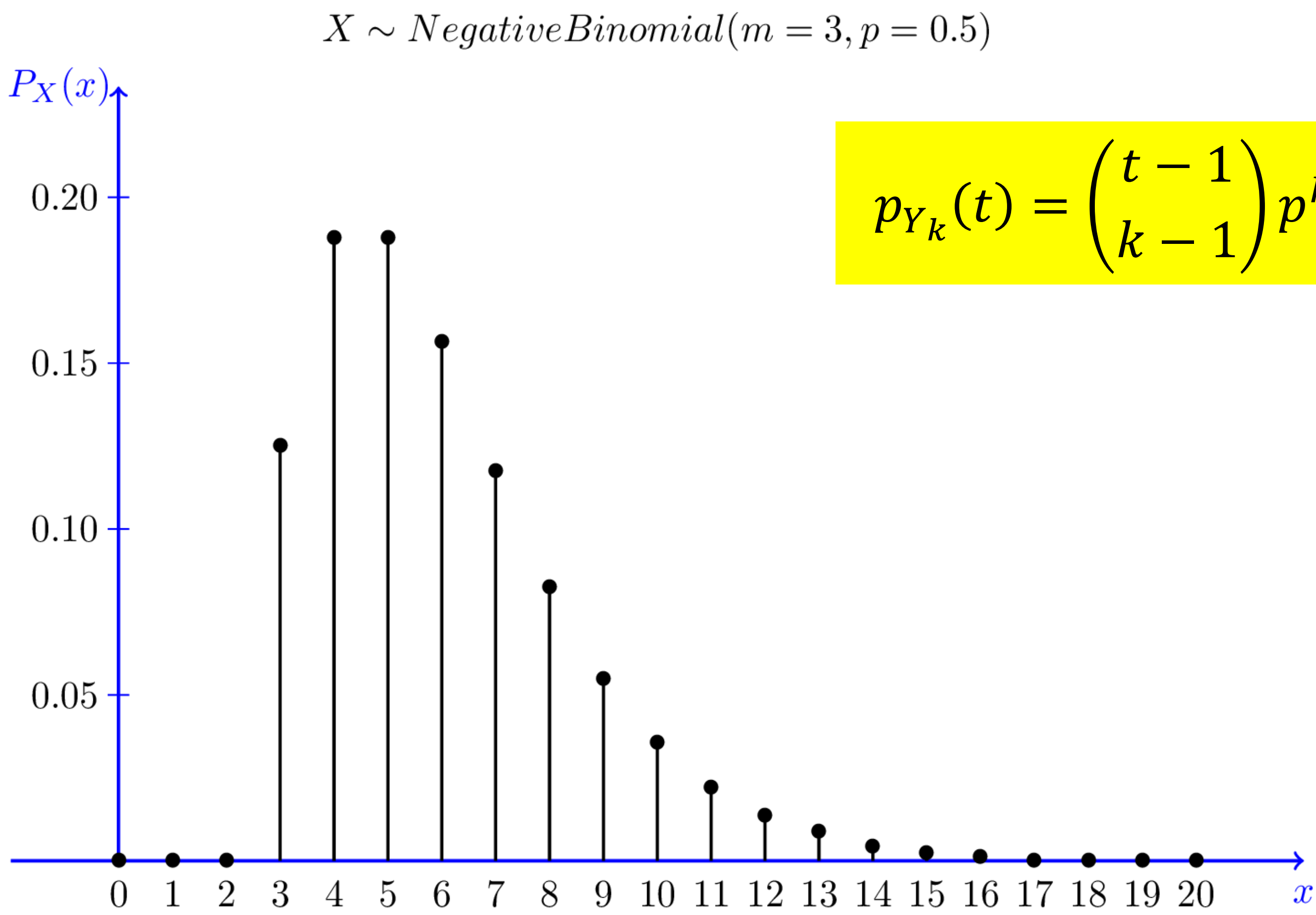
$$\Pr\{\text{Η } k\text{-οστή άφιξη γίνεται στο } t \text{ slot}\} = p_{Y_k}(t) =$$

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

$$= \boxed{\boldsymbol{\binom{t-1}{k-1} p^{k} (1-p)^{t-k}}}$$ **Negative Binomial**

$$\mathrm{E}[Y_k] = \mathbf{k/p} \qquad Var(Y_k) = \boldsymbol{k(1-p)/p^2}$$

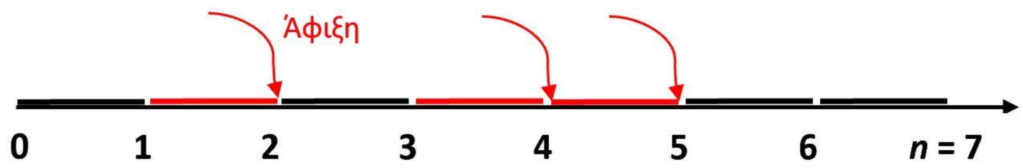
pmf της Negative Binomial



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

Σχετικές Τυχαίες Μεταβλητές (4)

- Η τ.μ. B που μετρά τον αριθμό των διαδοχικών slots στα οποία είχαμε αφίξεις (busy slots):



$\Pr\{k$ διαδοχικά busy slots$\} = p_B(k) =$

$= \boldsymbol{p^k(1-p)}$ **Γεωμετρική**

$\mathrm{E[B]} = \mathbf{1/(1-p)}$

$Var(\mathrm{B}) = \boldsymbol{p/(1-p)^2}$

Διαχωρισμός και Συνένωση Διαδικασιών Bernoulli

Διαχωρισμός

Οι διαδικασίες που προκύπτουν από το διαχωρισμό μίας διαδικασίας Bernoulli είναι επίσης Bernoulli.

$$
\begin{tikzcd}
BP(p) \arrow[r, "1-q"] \arrow[dr, "q"'] & BP((1-q)p) \\
 & BP(pq)
\end{tikzcd}
$$

Συνένωση

Η διαδικασία που προκύπτει από τη συνένωση δύο διαδικασιών Bernoulli είναι επίσης Bernoulli.

$$
\begin{tikzcd}
\boldsymbol{X_n} \quad BP(p) \arrow[r] & BP(p+q-pq) \\
\boldsymbol{Y_n} \quad BP(q) \arrow[ur] & \boldsymbol{Z_n}
\end{tikzcd}
$$

Συνένωση Διαδικασιών Bernoulli

- Θα έχουμε άφιξη της Z_n στο slot n, αν μία από τις X_n, Y_n ή και οι δύο ταυτόχρονα, έχουν άφιξη στο slot n.
- Δηλαδή: $Z_n = \max\{X_n, Y_n\}$

  και

  $$\boldsymbol{P}[\boldsymbol{Z_n} = \mathbf{1}] = 1 - P[X_n = 0, Y_n = 0] =$$
  $$= 1 - (1-p)(1-q) = \boldsymbol{p} + \boldsymbol{q} - \boldsymbol{pq}$$