Important Random Variables

CÁC BIẾN NGẪU NHIÊN (BNN) QUAN TRỌNG

Đặng Thanh Hải (Ph.D)
School of Engineering and Technology, VNUH
Homepage: uet.vnu.edu.vn/~hai.dang
Email: haidt82@yahoo.com

Acknowledgement

Hayder Radha
Associate Professor
Michigan State University
Department of Electrical & Computer Engineering

Examples of Discrete RVs

Các ví dụ về biến ngẫu nhiên (BNN) rời rạc

- **Bernoulli Random Variable**

  Biến ngẫu nhiên Bernoulli

  - Defined around a <u>single</u> event A

    Biến ngẫu nhiên Bernoulli

  - The outcome (of the underlying sample space S) is binary: the event A occurs <u>or</u> the event does not occur (i.e. A^c) $\Rightarrow$ Bernoulli trial

  - Transmit a packet: the packet is lost <u>or</u> it is not lost

    Truyền một gói tin: Gói tin có thể bị mất hoặc không

  - Transmit a bit: the bit is received correctly <u>or</u> a bit-error event occurs

    Truyền một bít: Bít có thể được nhận đúng hoặc có lỗi đã xẩy ra trên đường truyền

Kết quả (...) là nhị phân: Sự kiện A xẩy ra hoặc Sự kiện A không xẩy ra (...) => Phép thử Bernoulli

Examples of Discrete RVs

Các ví dụ về BNN rời rạc

- **Bernoulli Random Variable**

  Biến ngẫu nhiên Bernoulli

  - The Bernoulli random variable I is defined by mapping the event A into a real-number and its complement A^c into another real-number

  - Traditionally: $I(A)=1$; $I(A^c)=0$

    Thông thường:

  - $P[I=1] = p$; $P[I=0] = 1-p$

Biến ngẫu nhiên Bernoulli I được định nghĩa bằng cách ánh xạ sự kiện A vào một số thực và sự kiện đối ngẫu của nó Ac vào một số thực khác

$E[X] = p$ $\quad$ $\mathrm{VAR}[X] = p(1-p)$ Xem chứng minh ở mục 3.5.1 trong sách tiếng Anh

PhươngSai[X]

Examples of Discrete RVs

Ví dụ về BNN rời rạc

- **Binomial Random Variable**

  Biến ngẫu nhiên nhị thức

  - Defined based on a series of n <u>independent</u> Bernoulli trials

    Được định nghĩa dựa trên một dãy n phép thử Bernoulli độc lập

  - The binomial RV X is the number of occurrences of the event A in the n trials

    Biến ngẫu nhiên nhị thức X là số lần xẩy ra sự kiện A trong n phép thử

  Be careful about the "definition of the event A"

  Hãy cẩn thận về "Định nghĩa về sự kiện A"

  - Transmit n packets: count the number of packets that are <u>lost</u> Gửi n gói tin: Đếm số gói tin bị mất trên đường truyền
  - Transmit n bits: count the number of bits that are received <u>correctly</u>

    Gửi n bít: Đếm số bít được nhận đúng

Examples of Discrete RVs

Ví dụ về BNN rời rạc

- **Binomial Random Variable** Biến ngẫu nhiên nhị thức

  - If the Bernoulli random variable is defined by the mapping: $I_j(A)=1$ & $I_j(A^c)=0$ for $j=1, 2, \ldots, k, \ldots, n$ then the binomial RV X can be expressed:

Nếu biến ngẫu nhiên Bernoulli được định nghĩa bằng ánh xạ: ..... thì biến ngẫu nhiên nhị thức X có thể được biểu diễn:

$$X = \sum_{j=1}^{n} I_j$$

  - Therefore, $S_X=\{0, 1, 2, 3, \ldots, k, \ldots, n\}$

Do đó,

Examples of Discrete RVs

- **Binomial Random Variable**

  Biến ngẫu nhiên nhị thức

  - Let $P[I_j(A)=1] = p$, $P[I_j(A^c)=0] = 1-p \quad \forall\, j=1, 2, \ldots, n$

    Đặt

  - Then for the binomial RV: $X = \sum_{j=1}^{n} I_j$

    Do đó với biến ngẫu nhiên nhị thức:

  the probability that X=k , for k=0, 1, …, n:

  xác suất để X=k

$$\boxed{P[X=k] = \binom{n}{k} p^k (1-p)^{n-k}}$$

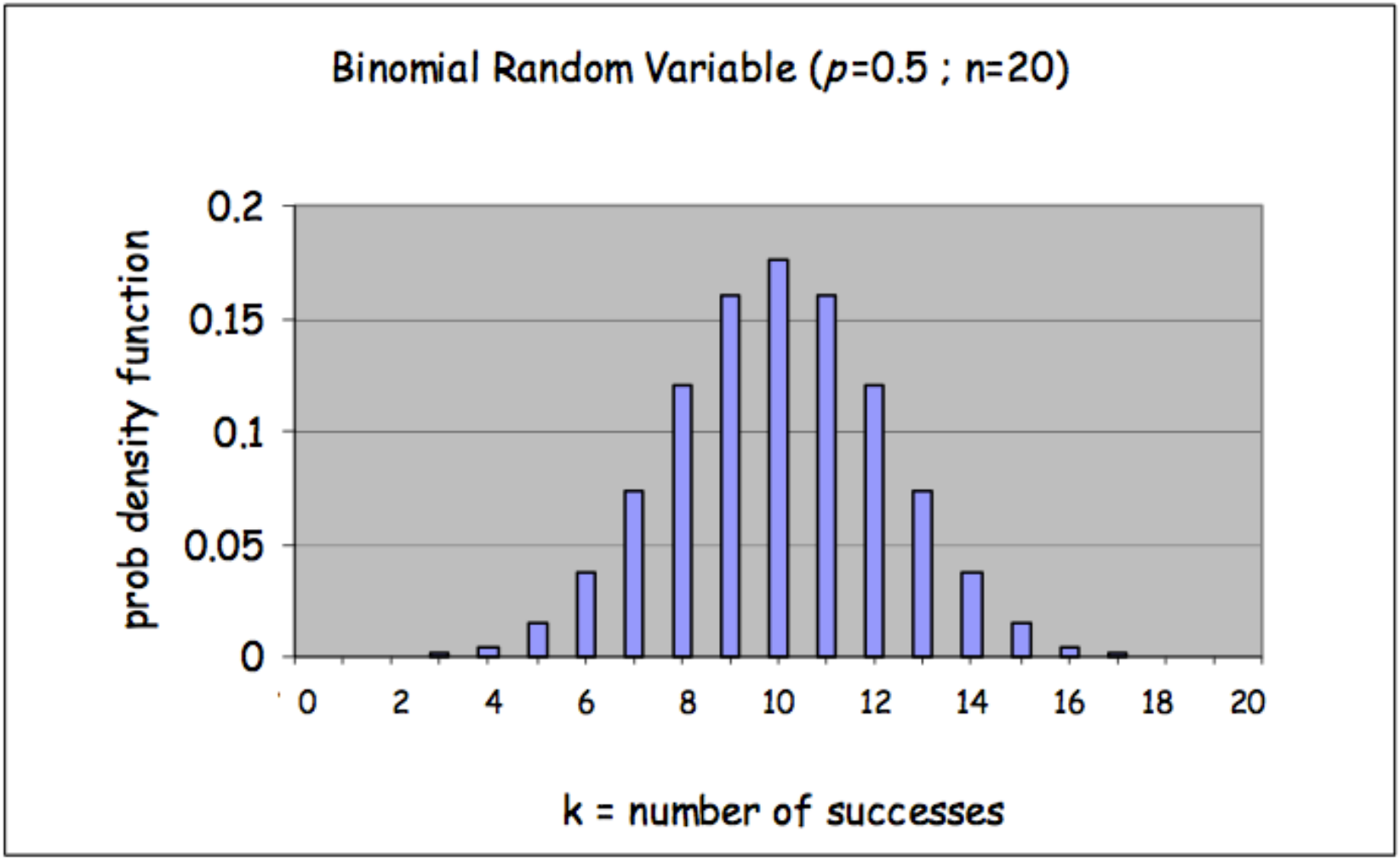
PhươngSai[X]

$E[X] = np \quad \mathrm{VAR}[X] = np(1-p)$

Xem chứng minh ở ví dụ 3.27 & 3.28 trong sách tiếng Anh

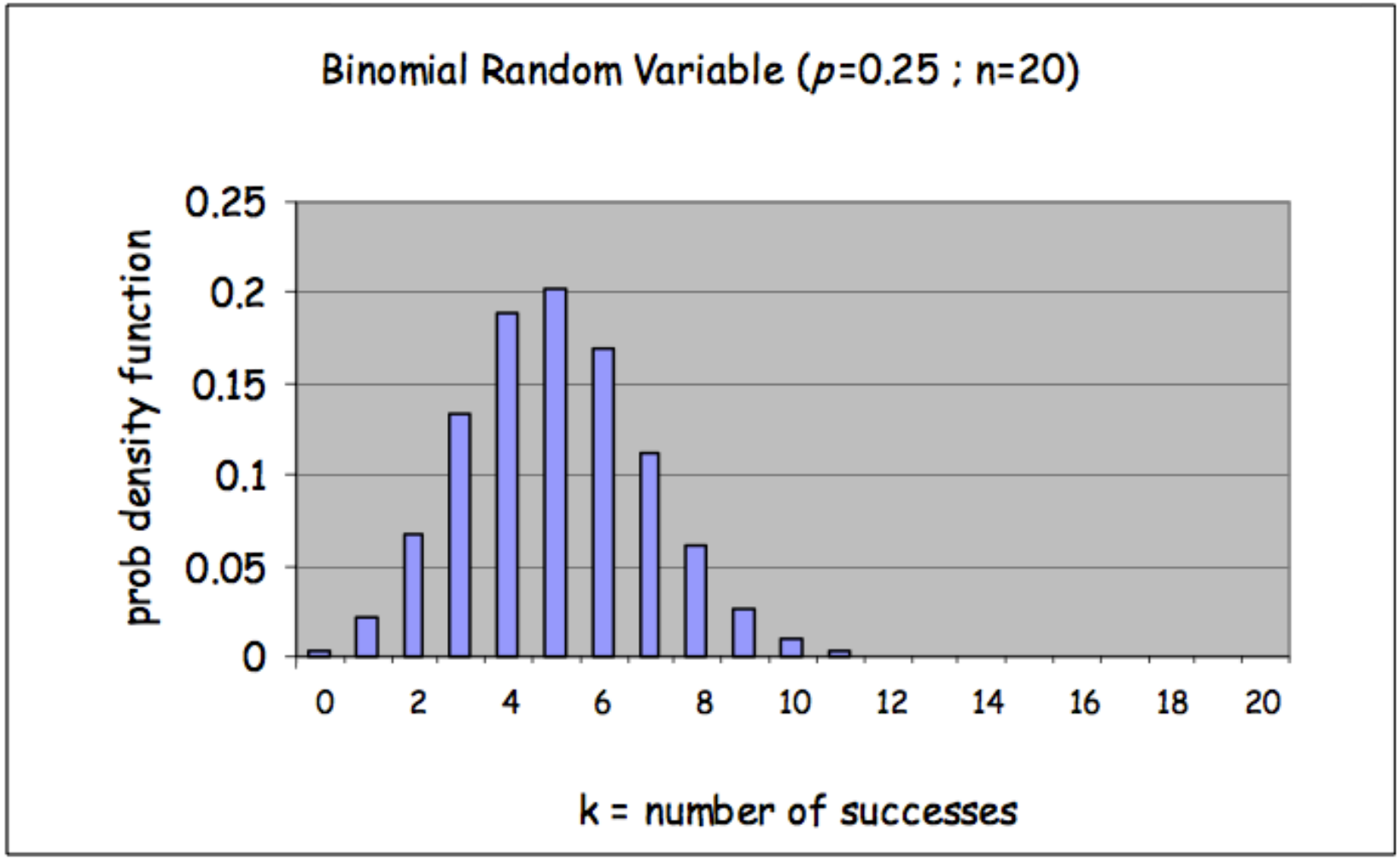
Binomial Random Variable

Biến ngẫu nhiên nhị thức

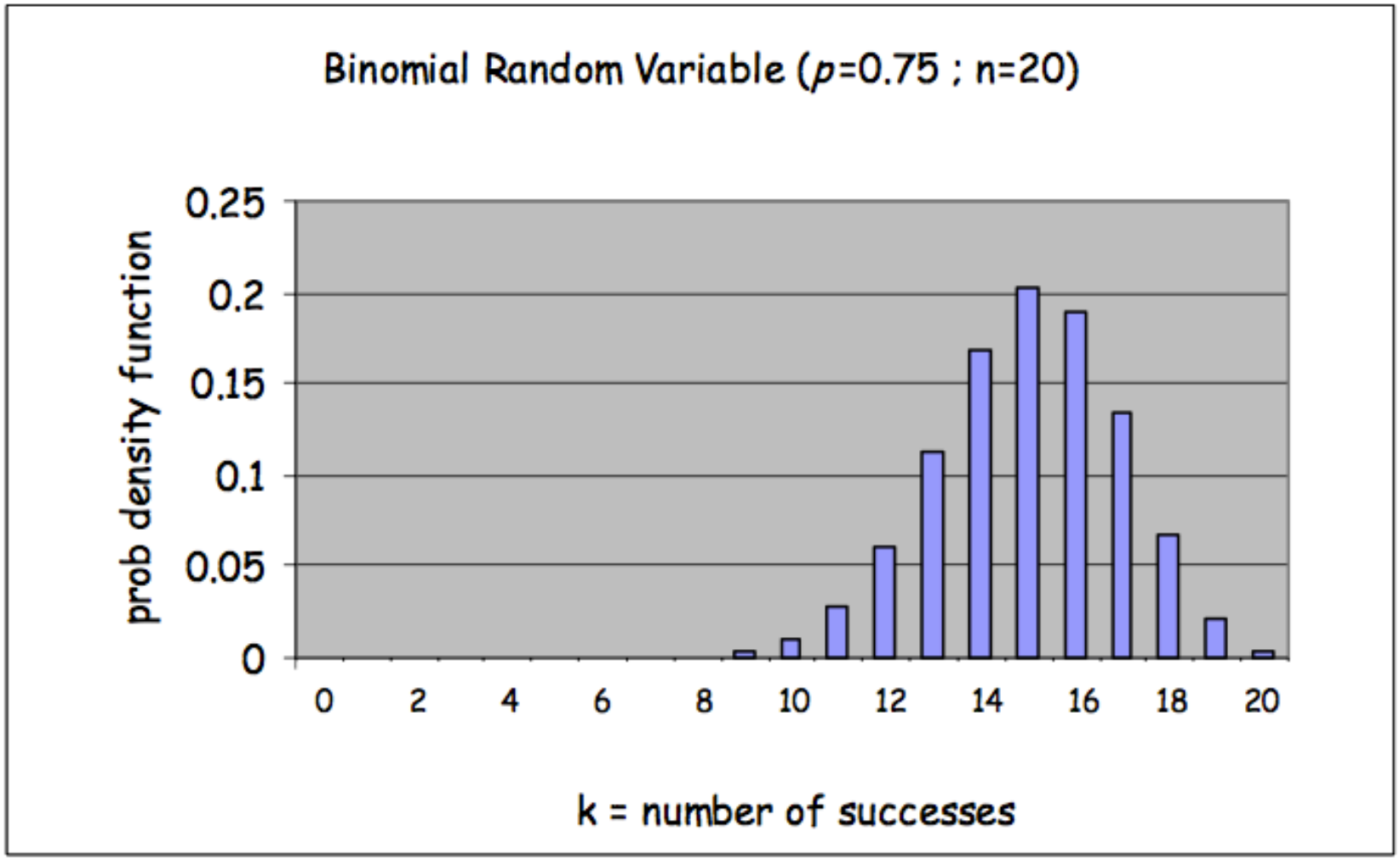


k = số lượng thành công

Binomial Random Variable

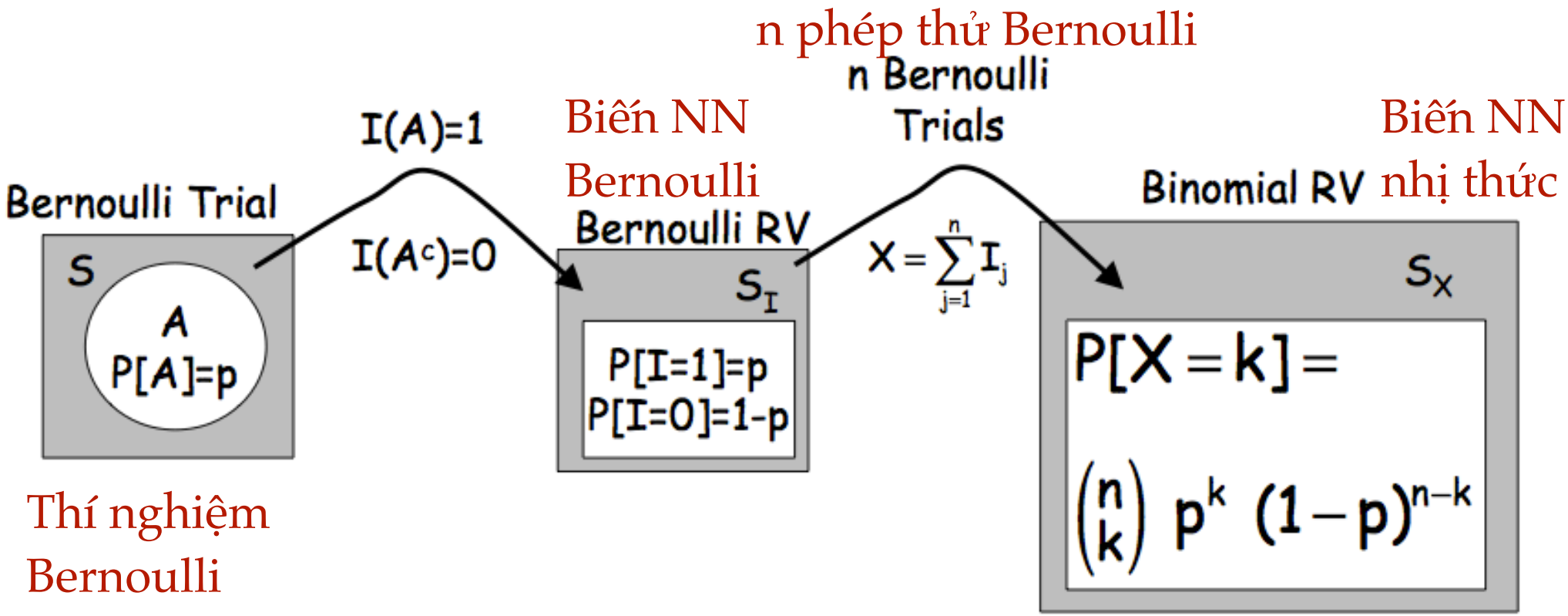


Binomial Random Variable



Bernoulli Trials and RVs

Thí nghiệm Bernoulli và các biến NN



Bernoulli Trials and RVs

Thí nghiệm Bernoulli và các biến NN

Example 3.29 Redundant Systems

A system uses triple redundancy for reliability: Three microprocessors are installed and the system is designed so that it operates as long as one microprocessor is still functional. Suppose that the probability that a microprocessor is still active after t seconds is $p = e^{-\lambda t}$. Find the probability that the system is still operating after t seconds.

Let X be the number of microprocessors that are functional at time t. X is a binomial random variable with parameter $n = 3$ and p. Therefore:

$$P[X \geq 1] = 1 - P[X = 0] = 1 - (1 - e^{-\lambda t})^3.$$

Ví dụ 3.29: Các hệ thống được thiết kế dư thừa

Để đảm bảo tính tin cậy, một hệ thống được thiết kế sử dụng 3 bộ vi xử lý sao cho nó có thể hoạt động được khi có một vi xử lý vẫn còn hoạt động được. Giả sử xác suất để một vi xử lý vẫn còn hoạt động được sau t giấy là p = ... Tìm xác suất để hệ thống này vẫn còn hoạt động được sau t giây.

Gọi X là số lượng vi xử lý vẫn còn hoạt động tài thời điểm t. X là một biến ngẫu nhiên nhị thức với tham số n=3 và p. Do đó

P[X>=1] = ...

Discrete RVs: Geometric

BNN rời rạc: Hình học/Nhân/Bội

- **Geometric random variable M**

  Biến ngẫu nhiên hình học/nhân/bội M

- **M is the number of Bernoulli trials <u>until</u> the first occurrence of the event A** M là số lượng phép thử Bernoulli cho đến khi xuất hiện lần đầu tiên sự kiện A

- **The number of: trials resulting in the event A^c ("failures") <u>plus</u> the first occurrence of A ("success")** Số lượng: phép thử sự kiện Ac ("thất bại") + sự xuất hiện lần đầu tiên của A ("thành công")

- **E.g., Sending a packet over the Internet: Keep re-transmitting until the packet makes it!**

  v.d., Gửi một gói tin trên mạng Internet: Tiếp tục gửi lại cho đến khi gói tin được xác nhận là đã được nhận thành công.

Discrete RVs: Geometric

Biến ngẫu nhiên rời rạc: Hình học/Nhân/Bội

- **Geometric random variable M has an infinitely countable possible values:**

Biến ngẫu nhiên hình học M nhận một tập đếm được vô hạn các giá trị

$$S_M=\{1, 2, 3, \dots k, \dots\}$$

Xác suất để M nhận giá trị k:

- **The probability that M takes on a value k:**

Lần "thành công" được tính vào khi đếm M

$$P[M=k] = (1-p)^{k-1}\, p$$

The "success" trial is included in counting M

$$E[M] = \frac{1-p}{p} \qquad \mathrm{VAR}[M] = \frac{1-p}{p^2}$$

- **Be careful, sometime M is defined such that:**

Để ý rằng, đôi lúc M có thể được định nghĩa là:

$$P[M=k] = (1-p)^{k}\, p$$

The "success" trial is not included in counting M

$$E[M] = \frac{1}{p} \qquad \mathrm{VAR}[M] = \frac{1-p}{p^2}$$

Lần "thành công" ko được tính vào khi đếm M

Chứng minh cách tính Kỳ vọng và phương sai: xem Ví dụ 3.15 và 3.22 trong sách Tiếng Anh)

Discrete RVs: Memory-less RV

Các BNN rời rạc: BNN không có bộ nhớ (**NÂNG CAO**)

- Given that after j trials the event A did not occur (i.e. j "failures"), what is the probability that we need at least other k trials <u>until</u> A occurs?

- If we have a "memory-less" experiment, the probability that "we need at least other k trials <u>until</u> we have a 'successful' trial" is independent of the past:

$$P[\{ M \geq k+j \} / \{M > j\}] = P[M \geq k]$$

Discrete RVs: Memory-less RV

Các biến ngẫu nhiên rời rạc: Biến ngẫu nhiên không có bộ nhớ (**NÂNG CAO**

- **The probability that a geometric RV M ≥ k:**

  Xác suất để một BNN hình học (bội) M >= k:

$$P[M \geq K] = \sum_{i=k}^{\infty} (1-p)^{i-1}\, p = (1-p)^{k-1}$$

- **Show that: a RV M satisfies the memory-less property if-and-only-if M is a geometric RV**

  Chứng minh rằng: Một BNN M là không nhớ
  nếu-và-chỉ-nếu M là một BNN hình học (bội)

Discrete RVs: Poisson

Biến ngẫu nhiên rời rạc: Poisson/Poát xông

- Poisson random variable N is the number of occurrences of a "completely random event" over a <u>unit of time</u>

Biến ngẫu nhiên Poisson N thể hiện số lần xuất hiện của một "sự kiện ngẫu nhiên" trong một đơn vị thời gian

- $S_N = \{0, 1, 2, 3, \ldots\}$

- E.g., the number of packets that arrive at a network router over one second; number of customers arriving at a queue between noon and 1pm; etc.

V.d., số lượng gói tin đến một router mạng trong một giây; số lượng khách hàng đến cửa hàng từ giữa trưa cho đến 1 giờ chiều; v.v.

Discrete RVs: Poisson

Biến ngẫu nhiên rời rạc: Poisson/Poát xông

- A Poisson RV is characterized by a parameter λ: Một biến ngẫu nhiên Poisson/Poát xông được đặc trưng bởi tham số ...:

  λ=the <u>average</u> number of occurrences per a unit of time

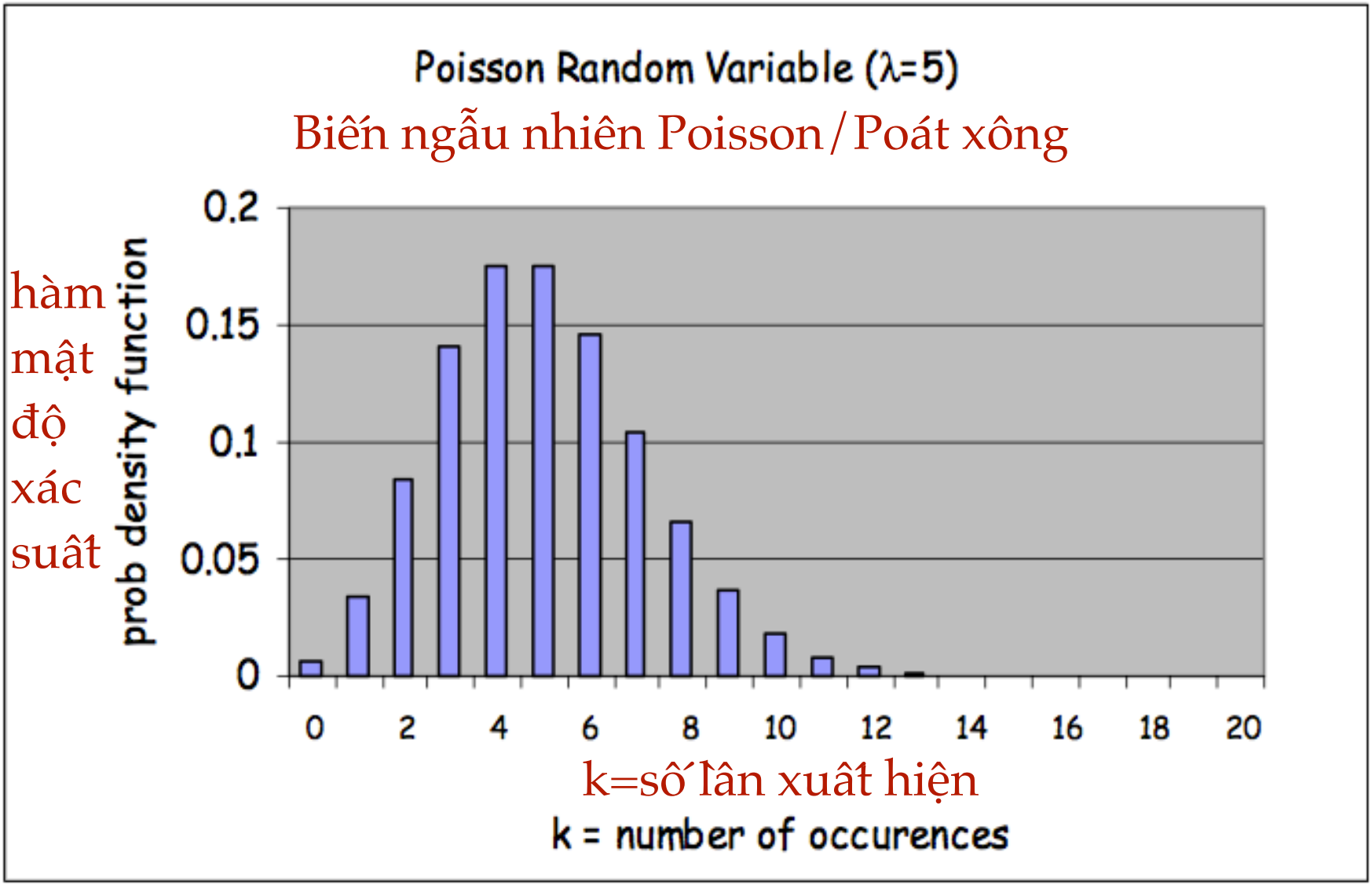
  ...=số lượng trung bình số lần xuất hiện trong một đơn vị thời gian

- The probability that a Poisson RV N takes on a value k (over a unit of time):

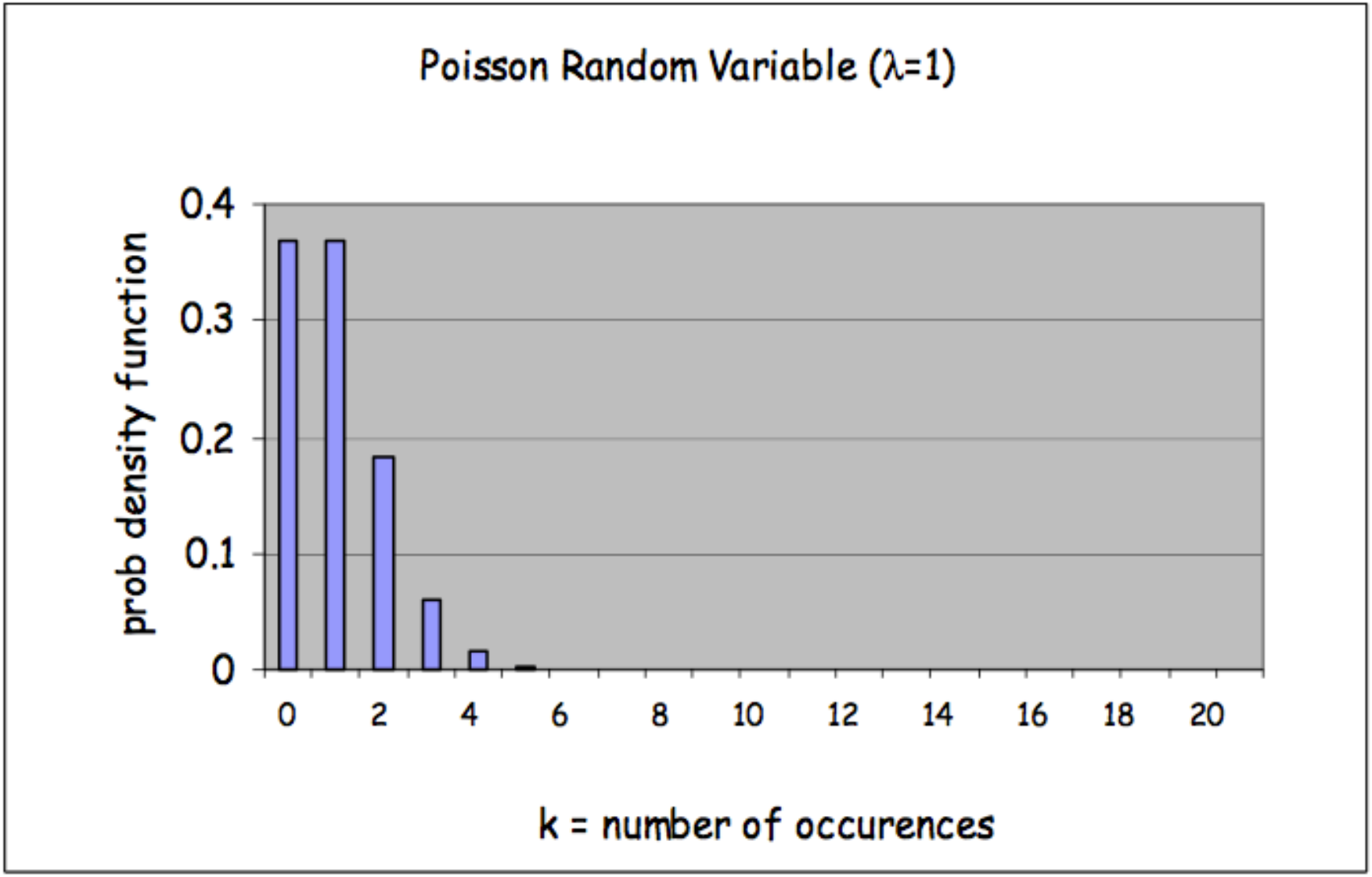
  Xác suất để một biến ngẫu nhiên Poisson nhận giá trị k (trong một đơn vị thời gian)

$$P[N=k]=\frac{\lambda^k}{k!}e^{-\lambda}$$

Discrete RVs: Poisson



Discrete RVs: Poisson



Discrete RVs: Poisson Moments

Biến ngẫu nhiên rời rạc: Mô men Poisson/Poát xông

- **The expected value (mean) of a Poisson RV:**

  Giá trị kỳ vọng (trung bình) của biến NN Poisson/Poát xông

$$E[N] = \sum_{k=0}^{\infty} k\, p_k = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = ?$$

- **The second moment of a Poisson RV:**

  Mô men bậc 2 của biến NN Poisson/Poát xông

$$E[N^2] = \sum_{k=0}^{\infty} k^2\, p_k = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = ?$$

$$E[X] = \alpha \qquad \mathrm{VAR}[X] = \alpha$$

Discrete RVs: Poisson Moments

BNN rời rạc: Các mô men Poisson/Poát xông

- To derive the moments of the Poisson random variable, you need to remember some old results from calculus.

  Để tính được mô men của biến ngẫu nhiên Poisson/Poát xông, bạn cần nhớ lại một số kiến thức giải tích kinh điển

  The infinite series expansion of e^{λ}:

  Khai triển chuỗi vô hạn của ...:

$$e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

Discrete RVs: Poisson

Các BNN rời rạc: Poisson/Poát xông

- λ=the <u>average</u> number of occurrences per <u>one unit of time</u> = Trung bình số lần xẩy ra trong một đơn vị thời gian
- The probability that a Poisson RV N takes on a value k over t units of time:

Xác suất để một BNN Poisson N nhận một giá trị k trong t đơn vị thời gian:

$$P[N = k \text{ in } t \text{ time units}] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

N= k trong t đơn vị thời gian

Discrete RVs: Poisson

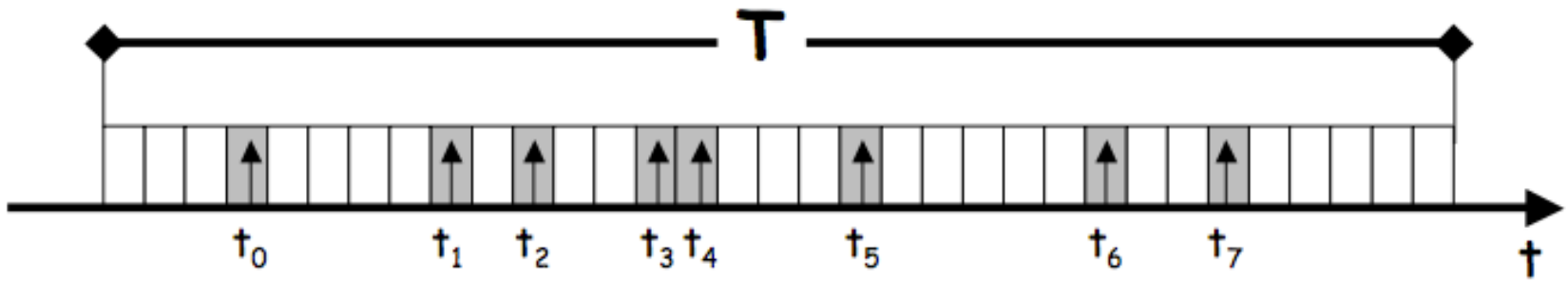
Nếu chúng ta chia một đơn vị thời gian Poisson T thành một số lượng rất lớn (v.d. n) khoảng thời gian rất ngắn ("tích tắc") để đảm bảo rằng:

- **If we divide a "Poisson time unit" T into a very-large number (e.g. n) of very-small time intervals ("ticks") such that:**
  - The event A can occur "at most" once in any "tick"

Sự kiện A chỉ có thể xẩy ra "nhiều nhất 1 lần" trong bất kỳ "tích tắc" nào

- **This can be viewed as n independent Bernoulli trials with P[A]=p=λ/n**

Điều này có thể xem như là n phép thử Bernoulli độc lập với P[A] = ...

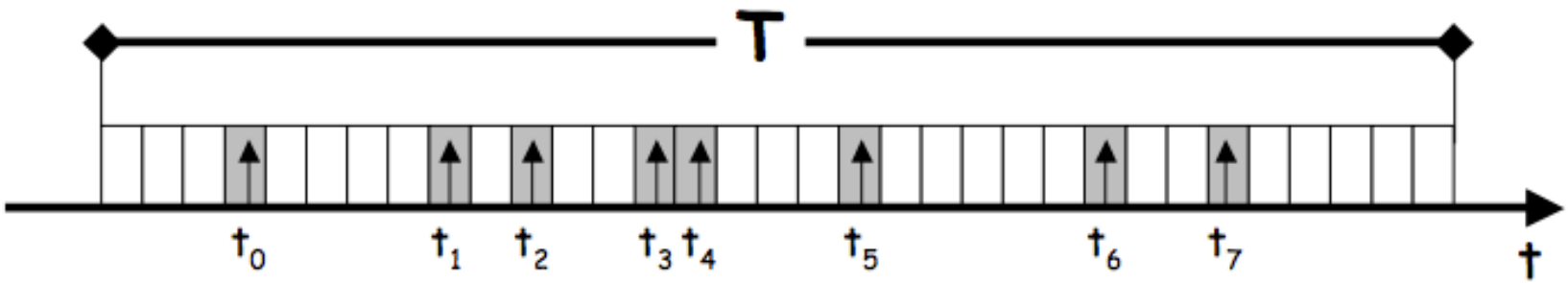


Discrete RVs: Poisson

Xác suất để sự kiện A xẩy ra k lần trong n phép thử này có thể được tính sử dụng hàm phân bố nhị thức (pmf) (với p = ...):

- **The probability that the event A occurs k times in these n trials can be computed using the binomial pmf (with p=λ/n):**

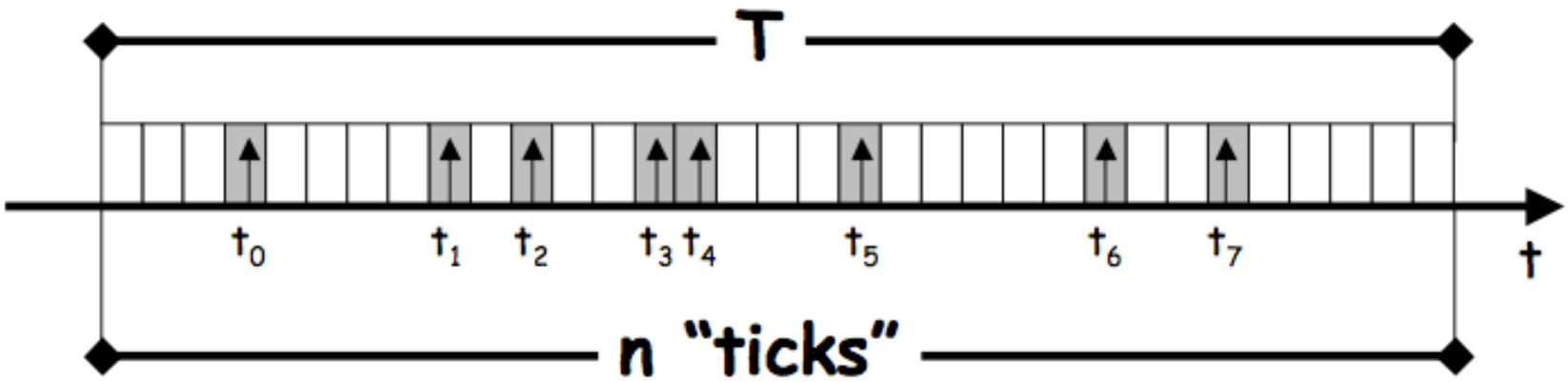
$$P[X=k]=\binom{n}{k}p^k(1-p)^{n-k}$$



Discrete RVs: Poisson

- This leads to the following approximation of the binomial pmf (with p=λ/n) by the Poisson pmf: Nó dẫn đến xấp xỉ sau cho hàm phân bố nhị thức pmf (với p=...) bằng hàm phân bố pmf Poisson

$$\boxed{\lim_{n\to\infty}\binom{n}{k}(\lambda/n)^k\,(1-(\lambda/n))^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda}}$$



n "tích tắc"

BNN Poisson: Ví dụ

Example 3.30 Queries at a Call Center

The number N of queries arriving in t seconds at a call center is a Poisson random variable with $\alpha = \lambda t$ where λ is the average arrival rate in queries/second. Assume that the arrival rate is four queries per minute. Find the probability of the following events: (a) more than 4 queries in 10 seconds; (b) fewer than 5 queries in 2 minutes.

The arrival rate in queries/second is $\lambda = 4$ queries/60 sec $= 1/15$ queries/sec. In part a, the time interval is 10 seconds, so we have a Poisson random variable with $\alpha = (1/15 \text{ queries/sec}) * 10 \text{ seconds} = 10/15$ queries. The probability of interest is evaluated numerically:

$$P[N > 4] = 1 - P[N \leq 4] = 1 - \sum_{k=0}^{4} \frac{(2/3)^k}{k!} e^{-2/3} = 6.33(10^{-4}).$$

In part b, the time interval of interest is $t = 120$ seconds, so $\alpha = 1/15 * 120 \text{ seconds} = 8$. The probability of interest is:

$$P[N \leq 5] = \sum_{k=0}^{5} \frac{(8)^k}{k!} e^{-8} = 0.10.$$

BNN Poisson: Ví dụ

Example 3.31 Arrivals at a Packet Multiplexer

The number N of packet arrivals in t seconds at a multiplexer is a Poisson random variable with $\alpha = \lambda t$ where λ is the average arrival rate in packets/second. Find the probability that there are no packet arrivals in t seconds.

$$P[N = 0] = \frac{\alpha^0}{0!} e^{-\lambda t} = e^{-\lambda t}.$$

This equation has an interesting interpretation. Let Z be the time until the first packet arrival. Suppose we ask, "What is the probability that $X > t$, that is, the next arrival occurs t or more seconds later?" Note that $\{N = 0\}$ implies $\{Z > t\}$ and vice versa, so $P[Z > t] = e^{-\lambda t}$. The probability of no arrival decreases exponentially with t.

Note that we can also show that

$$P[N(t) \geq n] = 1 - P[N(t) < n] = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

BNN Poisson: Ví dụ

Example 3.32 Errors in Optical Transmission

An optical communication system transmits information at a rate of 10^9 bits/second. The probability of a bit error in the optical communication system is 10^{-9}. Find the probability of five or more errors in 1 second.

Each bit transmission corresponds to a Bernoulli trial with a "success" corresponding to a bit error in transmission. The probability of k errors in $n = 10^9$ transmissions (1 second) is then given by the binomial probability with $n = 10^9$ and $p = 10^{-9}$. The Poisson approximation uses $\alpha = np = 10^9(10^{-9}) = 1$. Thus

$$P[N \geq 5] = 1 - P[N < 5] = 1 - \sum_{k=0}^{4} \frac{\alpha^k}{k!} e^{-\alpha}$$

$$= 1 - e^{-1}\left\{1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!}\right\} = .00366.$$

Discrete RVs: Negative Binomial RV

Các BNN rời rạc: BNN nhị thức âm

$S_X = \{r, r+1, \ldots\}$ where r is a positive integer

$$p_k = \binom{k-1}{r-1} p^r (1-p)^{k-r} \qquad k = r, r+1, \ldots$$

$$E[X] = \frac{r}{p} \qquad \mathrm{VAR}[X] = \frac{r(1-p)}{p^2} \qquad G_X(z) = \left(\frac{pz}{1-qz}\right)^r$$

Remarks: X is the number of trials until the rth success in a sequence of independent Bernoulli trials.

Nhận xét: X chỉ số phép thử đã thực hiện cho đến khi phép thử thành công thứ r trong dãy các phép thử Bernoulli độc lập

Discrete RVs: Uniform RV

Các BNN rời rạc: BNN phân bố đều

$$S_X = \{1, 2, \ldots, L\}$$

$$p_k = \frac{1}{L} \qquad k = 1, 2, \ldots, L$$

$$E[X] = \frac{L+1}{2} \quad \text{VAR}[X] = \frac{L^2 - 1}{12} \qquad G_X(z) = \frac{z}{L}\frac{1 - z^L}{1 - z}$$

Discrete RVs: Zipf RV

$S_X = \{1, 2, \ldots, L\}$ where L is a positive integer

$$p_k = \frac{1}{c_L}\frac{1}{k} \qquad k = 1, 2, \ldots, L \text{ where } c_L \text{ is given by Eq. (3.45)}$$

$$E[X] = \frac{L}{c_L} \qquad \mathrm{VAR}[X] = \frac{L(L+1)}{2c_L} - \frac{L^2}{c_L^2}$$

Remarks: The Zipf random variable has the property that a few outcomes occur frequently but most outcomes occur rarely.

Nhận xét: BNN Zipf có đặc tính là một số rất ít các kết quả (thí nghiệm) xẩy ra thường xuyên nhưng hầu hết lại rất hiếm khi xẩy ra.

Đc đặt theo tên của George Zipf người đã quan sát thấy rằng tần số xuất hiện của các từ trong một đoạn văn dài là tỉ lệ với hạng của chúng. Ban đầu nó được dùng rất nhiều để nghiên cứu sự phân bố của của cải, tài sản;

Gần đây là trong nhu cầu tính toán/ước lượng độ lớn của website, hành vi truy cập Web, và tính kết nối của Website. (Xem thêm ở mục 3.5.6 ở sách Tiếng Anh).

Examples of Random Variables

- Let us define the variable $\tau_n = t_{n+1} - t_n$

  Chúng ta quan tâm đến giá trị ... = tn+1 - tn

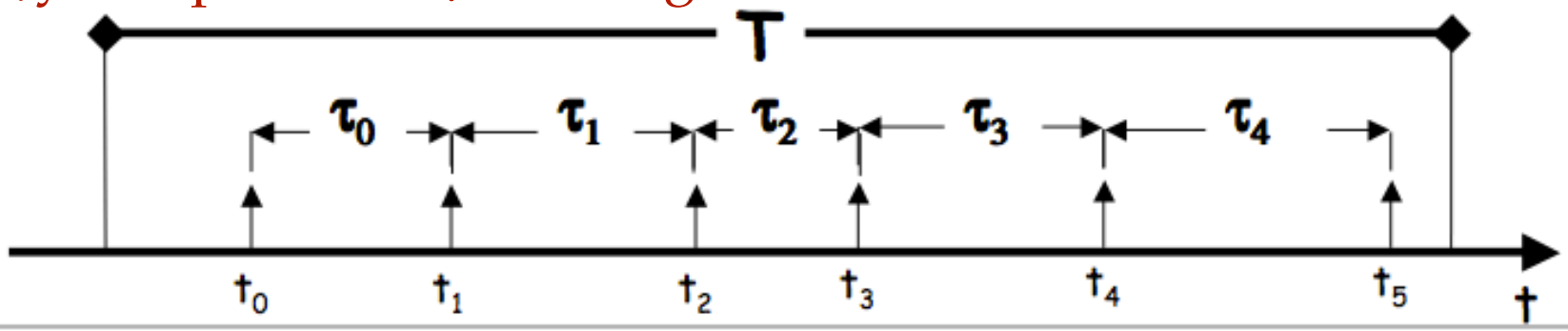
- Let $\tau = \{\tau_n\ ,\ n=0,\ 1,\ 2,\ \ldots.\}$

  Gọi ... = {..., n=0, 1, 2, ...}

- Therefore, τ measures the distance (in time) between the occurrences of a Poisson random event Khi đó, t đo khoảng cách (theo thời gian) giữa các lần xuất hiện của một sự kiện ngẫu nhiên Poisson

- Is τ a random variable?

  Vậy t có phải là một biến ngẫu nhiên?



Continuous RVs: Exponential

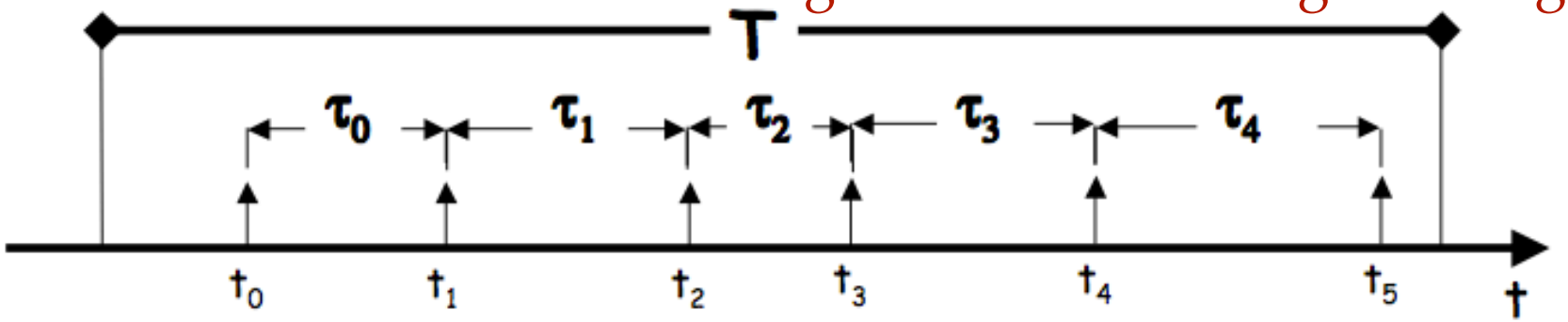
Các BNN liên tục: Mũ/Luỹ thừa

- Exponential RV X measures the inter-occurrence (inter-arrival) times among Poisson events

  BNN mũ/luỹ thừa X đo khoảng thời gian giữa các sự kiện Poisson

- E.g. Inter-arrival times of packets at a router; Inter-arrival times of customers at a queue in the bank

  V.d. Khoảng thời gian giữa các gói tin đến một router; khoảng thời gian giữa các khách hàng đến 1 ngân hàng



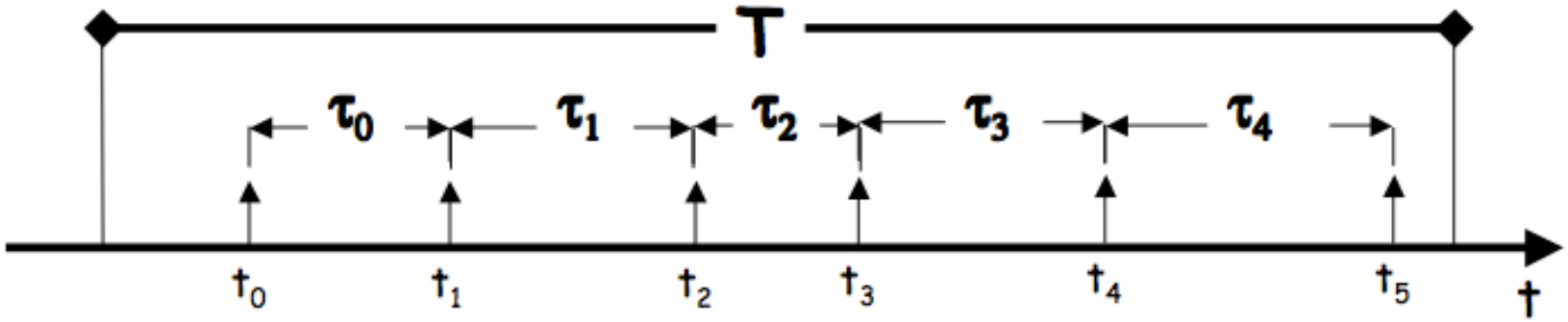
Continuous RVs: Exponential

BNN luỹ thừa X có thể được đặc trưng bởi tham số <lamda>.

- Exponential RV X can be characterized by the parameter λ which is the average # of arrivals per time unit T (same λ used in a Poisson random variable)

  <lamda> là số lần đến trung bình trong một đơn vị thời gian T (giống như trong BNN Poisson)
- ⇒ (1/λ) is average inter-arrival time

=> (1/<lamda>) là khoảng thời gian trung bình giữa các lần đến/xẩy ra



Continuous RVs: Exponential

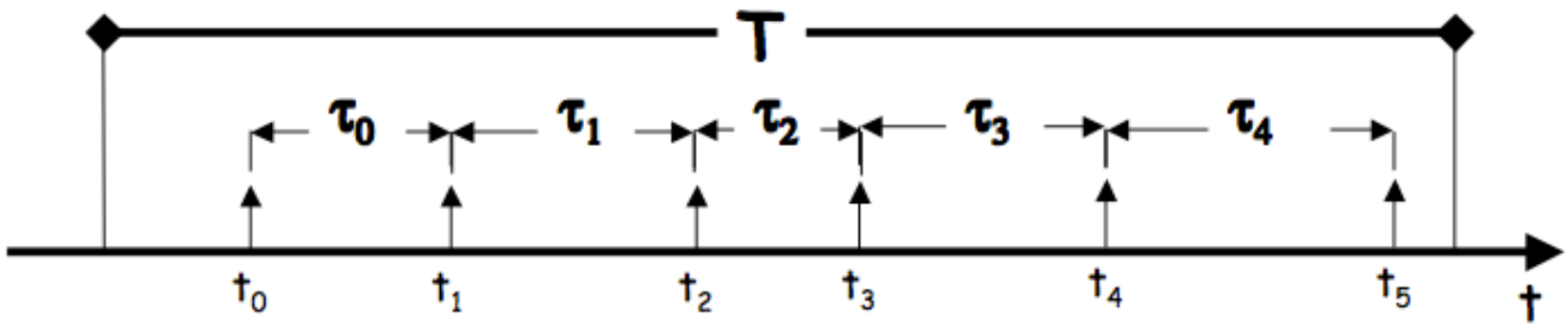
Các BNN liên tục: Mũ/Luỹ thừa

- The probability density function $f_X(\tau)$ of an exponential RV X is: Hàm mật độ xác suất fX(t) của một BNN luỹ thừa X là:

$$f_X(\tau) = \lambda e^{-\lambda\tau} \qquad \text{for} \quad \tau \geq 0 \quad \lambda > 0$$

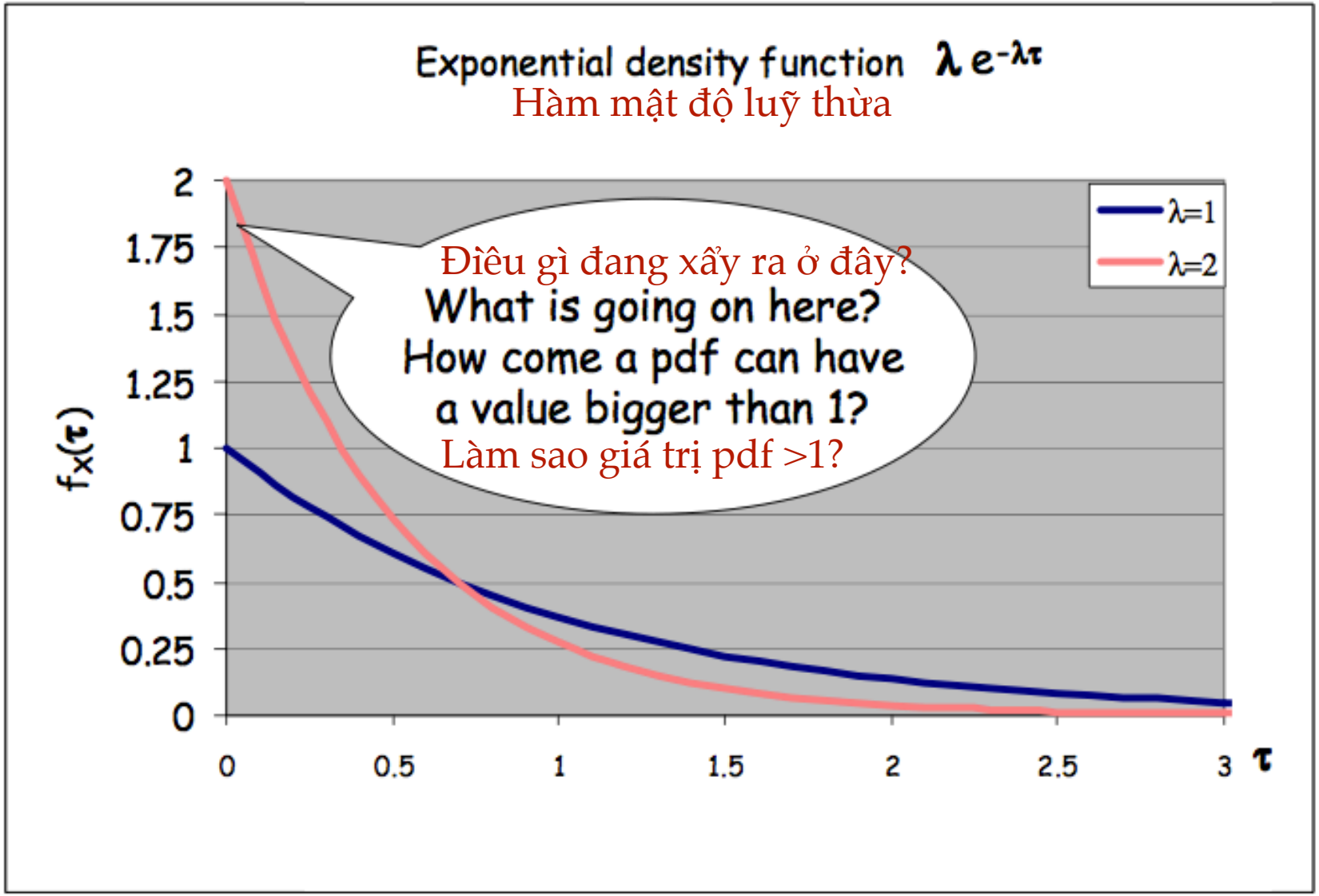
- where $(1/\lambda)$ is average inter-arrival time

trong đó (1/..) là khoảng thời gian trung bình giữa các lần đến/xẩy ra



Continuous RVs: Exponential

Các BNN liên tục: Mũ/Luỹ thừa



Continuous RVs: Exponential

Các BNN liên tục: Mũ / Luỹ thừa

Xác suất để một BNN luỹ thừa X nhận giá trị <= t?

- What is the probability that an exponential RV X takes on values less than or equal to τ?

- What is $F_X(\tau)$? Fx(t) là gì?

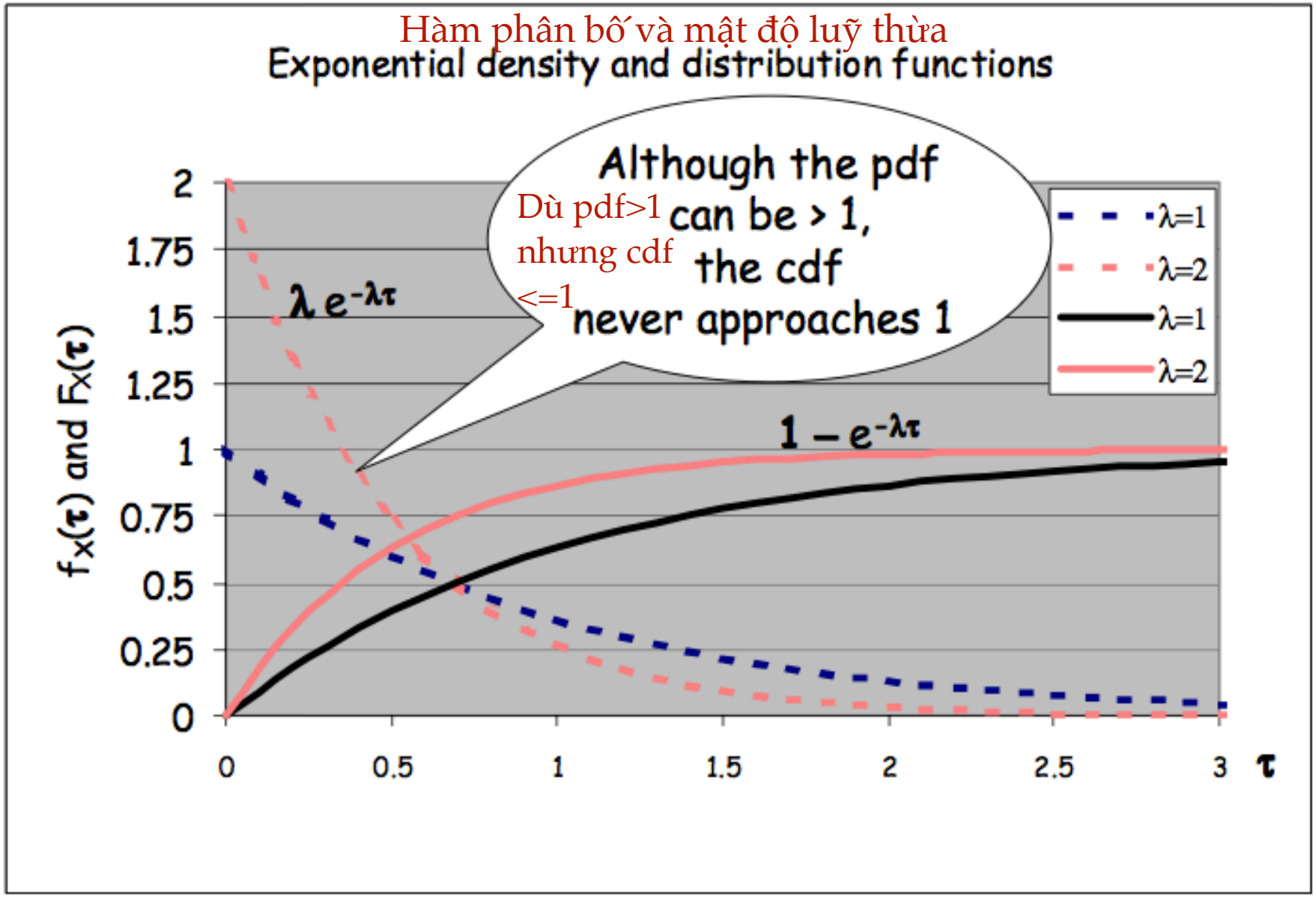
$$F_X(\tau) = \int_0^{\tau} f_X(t)dt = \int_0^{\tau} \lambda e^{-\lambda t} dt$$

$$\boxed{F_X(\tau) = 1 - e^{-\lambda\tau}}$$

Continuous RVs: Exponential

Các BNN liên tục: Mũ / Luỹ thừa

Hàm phân bố và mật độ luỹ thừa



Continuous RVs: Exponential

- **What is the expected value (mean) of an exponential RV X with a parameter λ?**

Kỳ vọng của một BNN luỹ thừa X với tham số <lamda> là gì?

$$m = \int_0^\infty t\, f_X(t)dt = \int_0^\infty t\, \lambda e^{-\lambda t} dt = \frac{1}{\lambda}.$$

Chứng minh: xem ví dụ 4.14 sách Tiếng Anh

Phương sai của một BNN luỹ thừa X với tham số <lamda> là gì?

- **What is the variance σ^2 of an exponential RV X with a parameter λ?**

$$\sigma^2 = \int_0^\infty (t-m)^2\, \lambda e^{-\lambda t} dt = \frac{1}{\lambda^2}$$

Continuous RVs: Memory-less RV

- Given that after waiting t seconds for the occurrence of an event A (e.g. arrival of a packet), what is the probability that we need to wait at least other h seconds until A occurs?

Tính không nhớ của BNN liên tục (**NÂNG CAO**)

- If we have a "memory-less" process, the probability that "we need to wait at least other h seconds until a packet arrives" is independent of the past:

$$P[\{X > t+h\} / \{X > t\}] = P[X > h]$$

Continuous RVs: Memory-less RV

- The exponential RV satisfies the memory-less property

  BNN luỹ thừa thoả mãn các đặc trưng không có bộ nhớ

- This can be shown by using:

Có thể chứng minh được rằng:

$$P[X > t] = e^{-\lambda t}$$

- Is there any relationship between the discrete memory-less RV (i.e. geometric RV) and the continuous memory-less RV (i.e. exponential)?

Có mối quan hệ giữa BNN không nhớ rời rạc (Biến hình học) và BNN không nhớ liên tục (Biến luỹ thừa)?

Continuous RVs: Memory-less RV

Đầu tiên, chia khoảng t-giây thành một số lượng lớn (k) các khoảng con

- First, let us divide a t-second interval into a large number (k) of sub-intervals

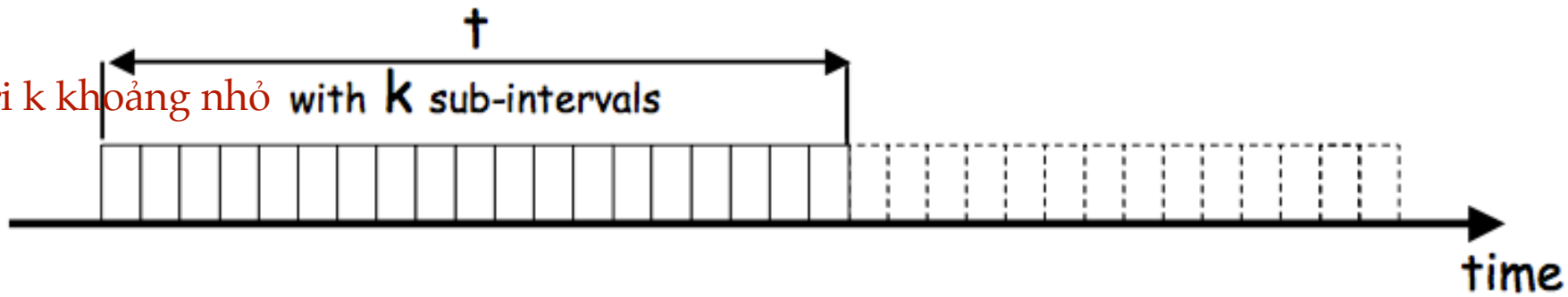
- Waiting t seconds with no occurrence of the event A ⇔ performing k trials with "failures"

Đợi t giây mà không thấy sự kiện A xẩy ra <=> thực hiện k phép thử và thất bại

Ukie cháu, tối chú sẽ mời Thảo, và cho cháu ké nhá :)))

$$P[X > t] = P[M > k]$$

với k khoảng nhỏ



Continuous RVs: Memory-less RV

- If M is a geometric RV: $P[M=k]=(1-p)^{k-1}p$

  Nếu M là một BNN hình học (bội): P[M=k]=....

$\Rightarrow P[X > t] = P[M > k] = (1-p)^k$

Bằng cách thay p=(.....) ....

By using $p = (\lambda t/k)$ & $(1-(a/k))^k \rightarrow e^{-a}$ as $k \rightarrow \infty$

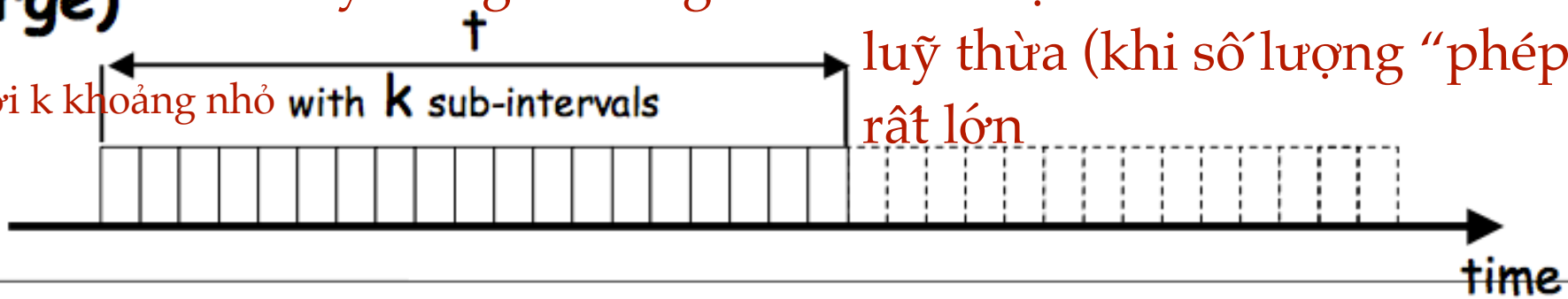
this leads to:

nó dẫn đến:

$$\lim_{k\to\infty} P[M>k] = \lim_{k\to\infty}\left(1-\frac{\lambda t}{k}\right)^k = e^{-\lambda t}$$

- This implies that a geometric RV leads to the exponential RV (when the number of "trials" is very large) Điều này có nghĩa rằng BNN hình học sẽ trở thành BNN luỹ thừa (khi số lượng "phép thử" rất lớn

với k khoảng nhỏ with k sub-intervals



The "Bernoulli Islands"

Phép thử Bernoulli

Bernoulli Trial

$S=\{A, A^c\}$

A

$P[A]=p$

$I(A)=1$

$I(A^c)=0$

BNN Bernoulli

Bernoulli RV

$S_I=\{0, 1\}$

$P[I=1]=p$

$P[I=0]=1-p$

n phép thử Bernoulli

n Bernoulli Trials

$X=\sum_{j=1}^{n} I_j$

BNN nhị thức

Binomial RV

$S_X=\{0, 1, \ldots, n\}$

$P[X=k]=\binom{n}{k} p^k (1-p)^{n-k}$

Count # of "failures" until a "success" occurs

Đếm số lần thất bại cho đến lần thành công đầu tiên

<lamda> là trung bình số lần "thành công" trong 1 đơn vị thời gian T; p = (....)

λ is average # of "successes" over a time T; $p=(\lambda/n)$

Số phép thử lớn n -> <vôcùng> trên đơn vị thời gian T

Không nhớ

Memory-less

BNN luỹ thừa

Exponential RV

$S_X=\{t: t\geq 0\}$

$f_X(t) = \lambda e^{-\lambda t}$

Geometric RV

$S_M=\{1, 2, \ldots, \}$

$P[M=k]=(1-p)^{k-1}p$

Large # of "failures" $k \to \infty$ over a time t

Số lần "thất bại" lớn k -> <vôcùng> trên 1 đơn vị thời gian t

Measure inter-arrival time of Poisson events

Đo khoảng thời gian giữa các sự kiện Poisson

Poisson RV

BNN Poisson

$S_N=\{0, 1, \ldots, \}$

$P[N=k] = \frac{\lambda^k}{k!} e^{-\lambda}$

Continuous RVs: Gaussian

Các BNN liên tục: Gaussian

- **Gaussian RV X with mean m and variance σ^2 has a probability density function:**

  BNN Gaussian (chuẩn) X với kỳ vọng m và phương sai ... có hàm mật độ xác suất:

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma^2}$$

- **The cdf:** Hàm phân bố xác suất tích luỹ (cdf)

$$F_x(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-(s-m)^2/2\sigma^2} ds$$

The Q-function

Hàm Q

$$P[X \le x] = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-(x'-m)^2/2\sigma^2}\,dx'.$$

The change of variable $t = (x' - m)/\sigma$ results in

Đặt biến t = (x'-m)/... sẽ dẫn đến

$$F_X(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(x-m)/\sigma} e^{-t^2/2}\,dt$$

$$= \Phi\left(\frac{x-m}{\sigma}\right)$$

where $\Phi(x)$ is the cdf of a Gaussian random variable with $m = 0$ and $\sigma = 1$:

trong đó ...(x) là hàm cdf của bnn (chuẩn) Gaussian với m=0 và ..=1:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt.$$

The Q-function

Hàm Q

$$Q(x) = 1 - \Phi(x)$$
$$= \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2}\, dt.$$

- Approximation required by two function as followed

Việc xấp xỉ yêu cầu hai hàm như sau:

$$Q(x) \simeq \left[\frac{1}{(1-a)x + a\sqrt{x^2+b}} \right] \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

where $a = 1/\pi$ and $b = 2\pi$ [Gallager]

trong đó a=... và b= [theo Gallager]

$$Q(x) = 10^{-k}$$

TABLE 4.2 Comparison of $Q(x)$ and approximation given by Eq. (4.54).

Bảng 4.2. So sánh Q(x) và giá trị xấp xỉ của nó theo phương trình 4.54.

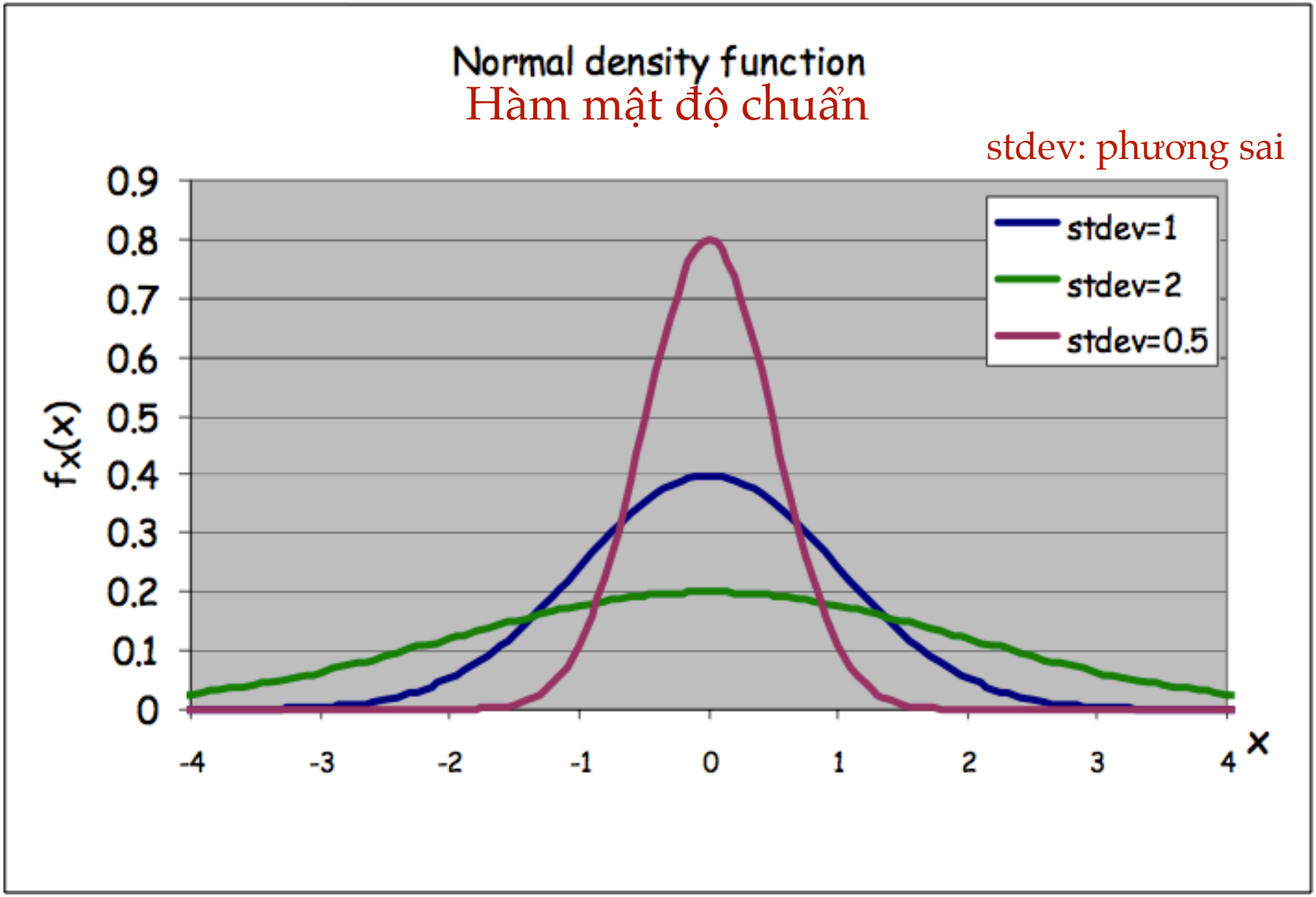
x	$Q(x)$	Approximation giá trị xấp xỉ	x	$Q(x)$	Approximation giá trị xấp xỉ
0	5.00E-01	5.00E-01	2.7	3.47E-03	3.46E-03
0.1	4.60E-01	4.58E-01	2.8	2.56E-03	2.55E-03
0.2	4.21E-01	4.17E-01	2.9	1.87E-03	1.86E-03
0.3	3.82E-01	3.78E-01	3.0	1.35E-03	1.35E-03
0.4	3.45E-01	3.41E-01	3.1	9.68E-04	9.66E-04
0.5	3.09E-01	3.05E-01	3.2	6.87E-04	6.86E-04
0.6	2.74E-01	2.71E-01	3.3	4.83E-04	4.83E-04
0.7	2.42E-01	2.39E-01	3.4	3.37E-04	3.36E-04
0.8	2.12E-01	2.09E-01	3.5	2.33E-04	2.32E-04
0.9	1.84E-01	1.82E-01	3.6	1.59E-04	1.59E-04
1.0	1.59E-01	1.57E-01	3.7	1.08E-04	1.08E-04
1.1	1.36E-01	1.34E-01	3.8	7.24E-05	7.23E-05
1.2	1.15E-01	1.14E-01	3.9	4.81E-05	4.81E-05
1.3	9.68E-02	9.60E-02	4.0	3.17E-05	3.16E-05
1.4	8.08E-02	8.01E-02	4.5	3.40E-06	3.40E-06
1.5	6.68E-02	6.63E-02	5.0	2.87E-07	2.87E-07
1.6	5.48E-02	5.44E-02	5.5	1.90E-08	1.90E-08

TABLE 4.3 $Q(x) = 10^{-k}$

k	$x = Q^{-1}(10^{-k})$
1	1.2815
2	2.3263
3	3.0902
4	3.7190
5	4.2649
6	4.7535
7	5.1993
8	5.6120
9	5.9978
10	6.3613

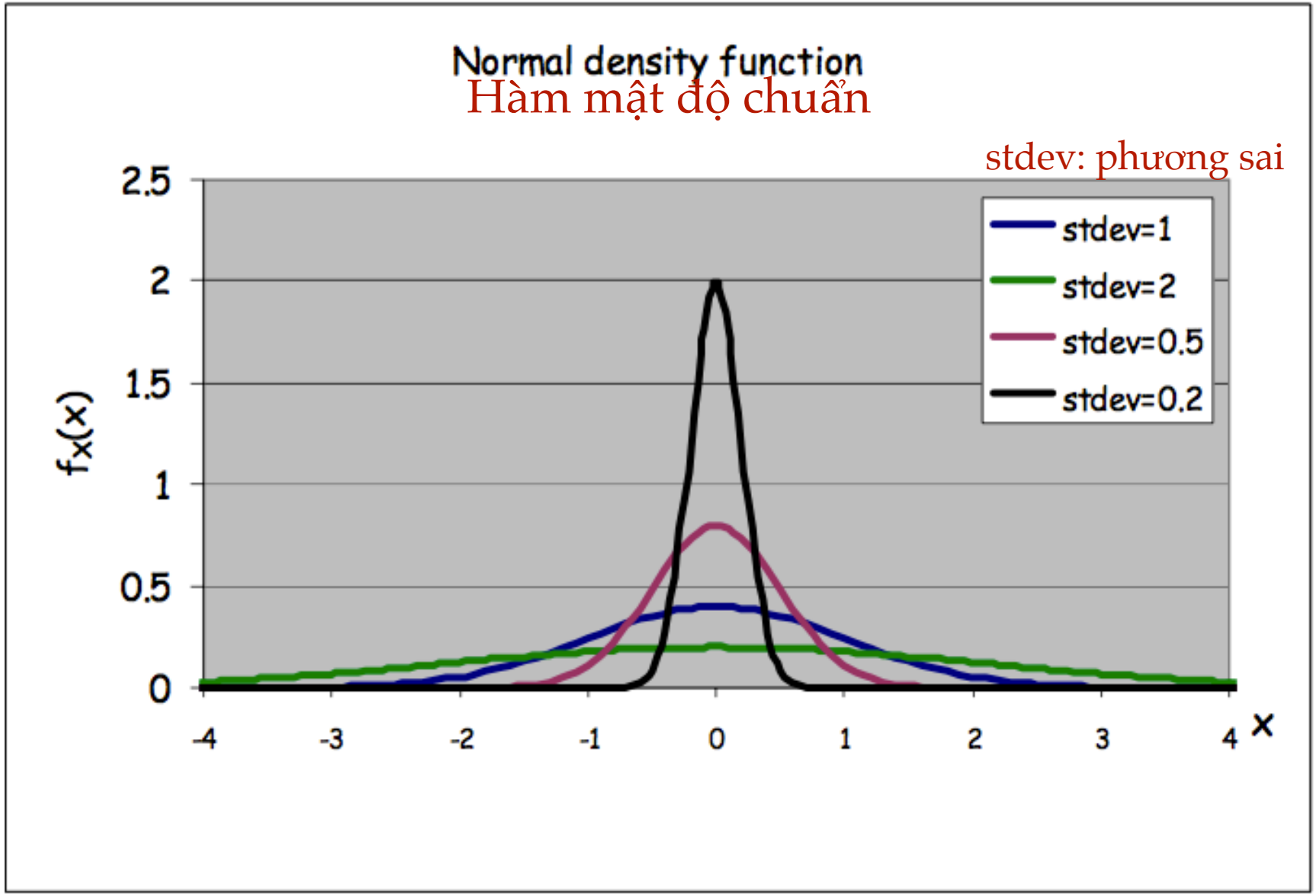
Continuous RVs: Gaussian (Normal)

Các BNN liên tục: Gaussian (Chuẩn)



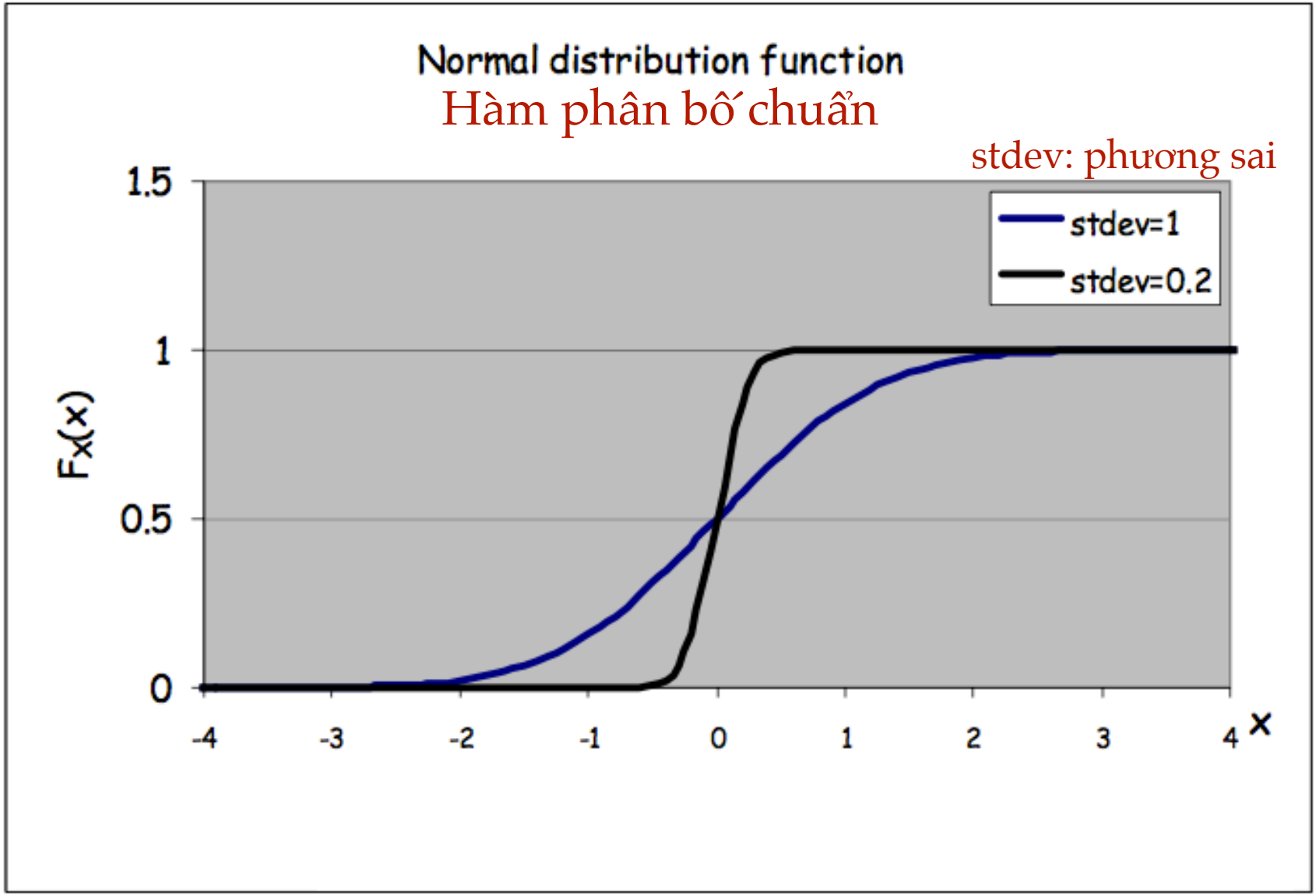
Continuous RVs: Gaussian (Normal)

Các BNN liên tục: Gaussian (Chuẩn)



Continuous RVs: Gaussian (Normal)

Các BNN liên tục: Gaussian (Chuẩn)



Laplacian Random Variable

BNN Laplacian

Probability density Function

Hàm mật độ xác suất

$$f_x(x) = \frac{\alpha}{2} e^{-\alpha|x|}$$

Phương sai

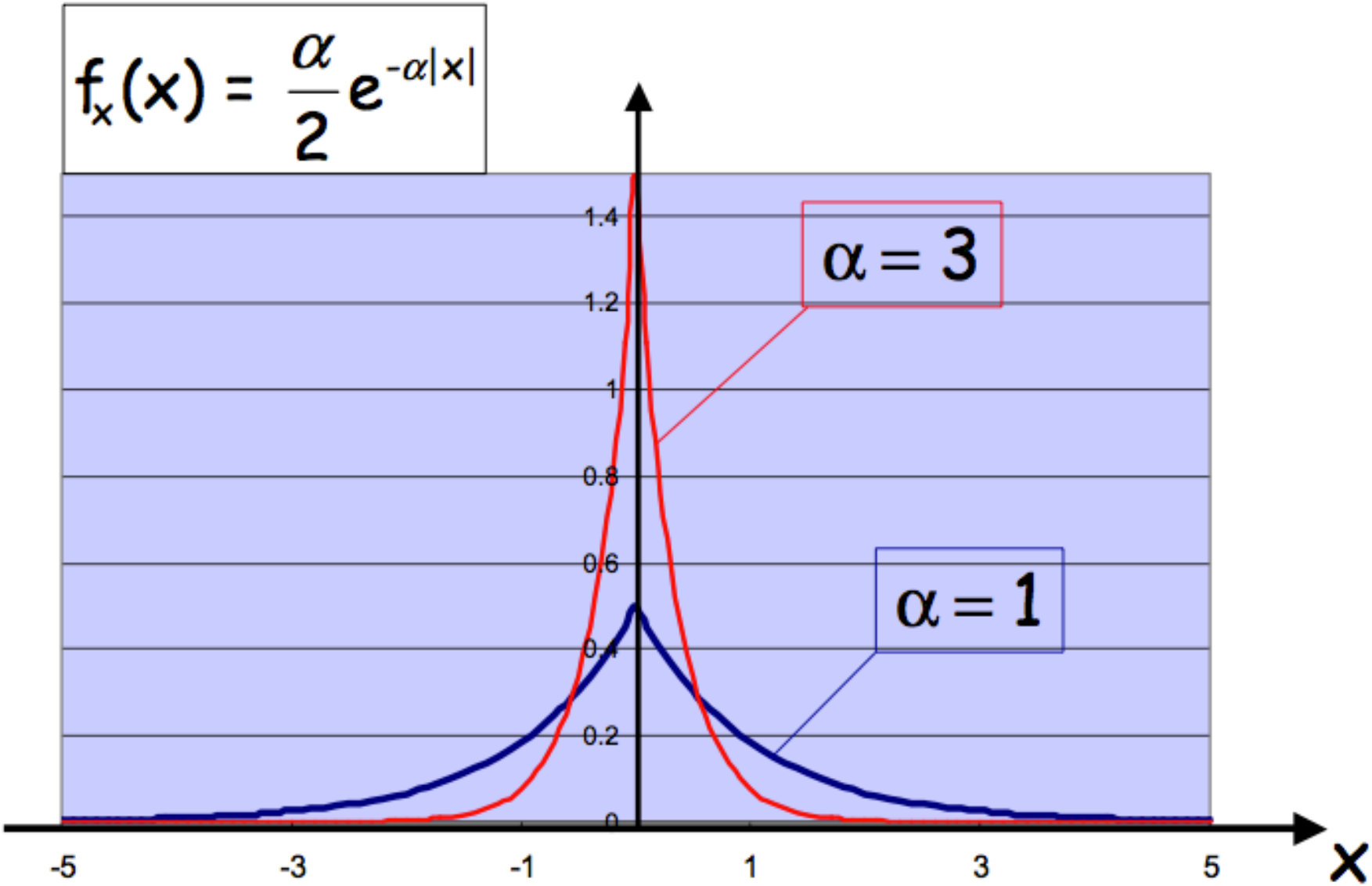
Variance = $2/\alpha^2$

- **The cdf:** Hàm phân bố tích luỹ cdf

$$F_x(x) = \frac{\alpha}{2} \int_{-\infty}^{x} e^{-\alpha|s|} ds$$

Laplacian Random Variable

BNN Laplacian



The Gamma Random Variable

BNN Gamma

- The Gamma Random Variable captures a wide range of random variables and is used in many applications: BNN Gamma tổng quá hoá được một lượng lớn các bnn và được sử dụng trong rât nhiều ứng dụng:
  - Queuing systems (e.g., multiple queues);
    Các hệ thống hàng đợi (v.d., hệ thống nhiều hàng đợi)
  - Reliability;
    Đảm bảo tính ổn định
  - VLSI chips' defects
    Lỗi của chíp VLSI

The Gamma Random Variable

BNN Gamma

Hàm mật độ pdf của bnn Gamma đc đặc trưng bởi 2 tham số:

- **The pdf function of the Gamma R.V. is characterized by two parameters: λ and α**
  - The "lambda" is the same λ used in the pdf function of the exponential R.V.

    Tham số "lamda" là tương tự như trong hàm mật độ pdf của bnn luỹ thừa

$$f_X(x) = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

$$0 < x < \infty$$
$$\lambda > 0$$
$$\alpha > 0$$

$$\Gamma(\alpha) = \int_0^{\infty} s^{\alpha-1} e^{-s} ds$$

Gamma Function

Hàm Gamma

The Gamma Random Variable

BNN Gamma

BNN Gamma

Gamma R.V.

$$f_X(x) = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

$$\alpha = 1$$

$$f_X(x) = \lambda e^{-\lambda x}$$

Exponential R.V.

BNN luỹ thừa

Các đặc trưng của hàm Gamma

Properties of the Gamma Function

$$\Gamma(1) = 1$$

$$\Gamma(n) = (n-1)!$$

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$$

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$$

The n-Erlang RV

BNN n-Erlang

BNN Gamma

Gamma R.V.

$$f_X(x) = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

$\alpha = n$

$= \text{positive integer}$

số nguyên dương

$$f_X(x) = \frac{\lambda^{n} x^{n-1} e^{-\lambda x}}{\Gamma(n)} = \frac{\lambda^{n} x^{n-1} e^{-\lambda x}}{(n-1)!}$$

"n-Erlang" R.V.

Các đặc trưng của hàm Gamma

Properties of the Gamma Function

$$\Gamma(1) = 1$$

$$\Gamma(n) = (n-1)!$$

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$$

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$$

Remarks: An m–1 Erlang random variable is obtained by adding m independent exponentially distribute random variables with parameter λ.

Chú ý: Một bnn Erlang m-1 sẽ đc tạo ra bằng cách cộng m bnn luỹ thừa có tham số <lamda> độc lập nhau

The Gamma RV

BNN Gamma

BNN Chi-bình phương với k bậc tự do: ...., k là một số nguyên dương và

Chi-Square Random Variable with k degrees of freedom: $\alpha = k/2$, k a positive integer, and $\lambda = 1/2$

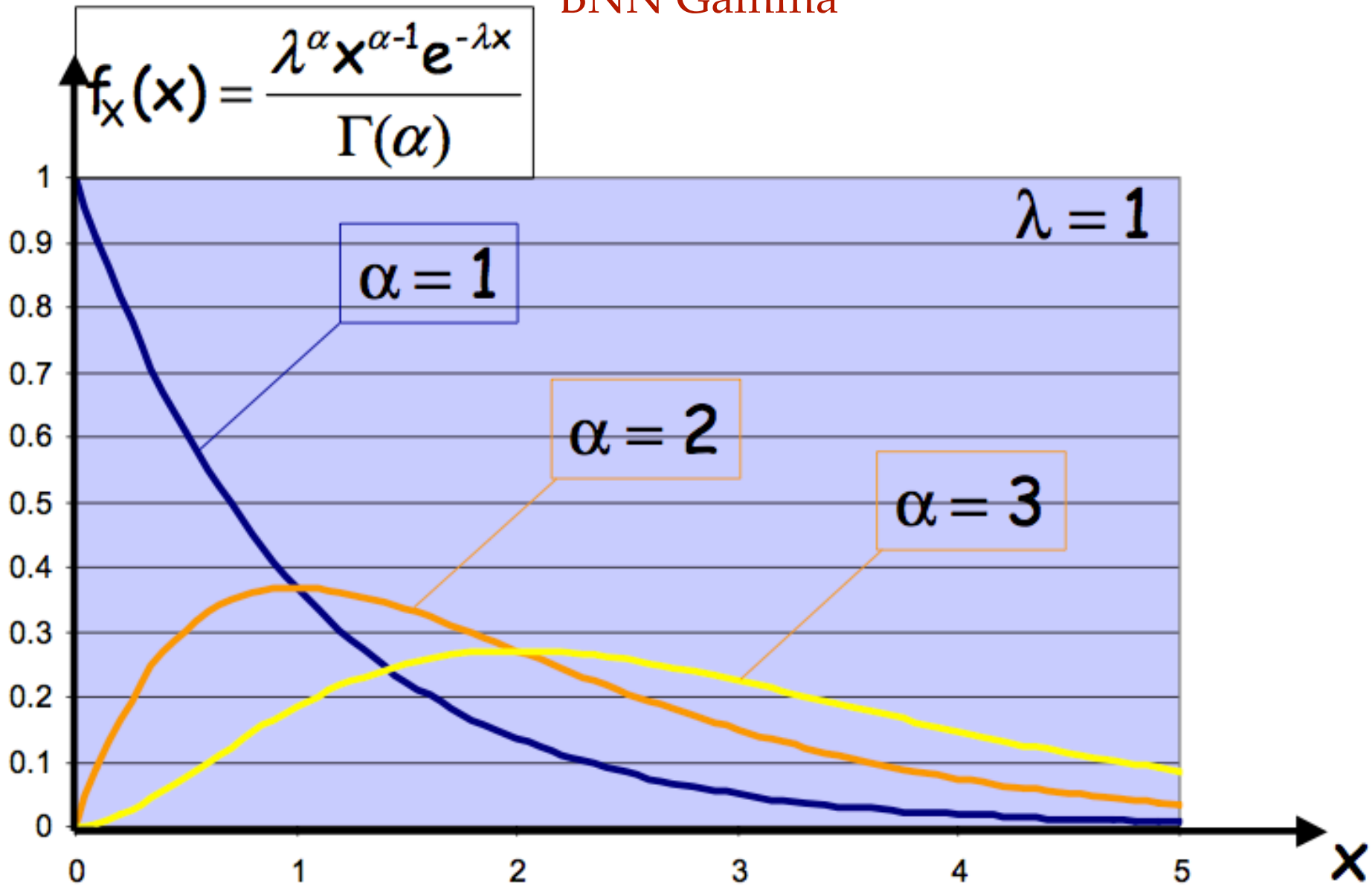
$$f_X(x) = \frac{x^{(k-2)/2} e^{-x/2}}{2^{k/2}\Gamma(k/2)} \qquad x > 0 \qquad \Phi_X(\omega) = \left(\frac{1}{1 - 2j\omega}\right)^{k/2}$$

Remarks: The sum of k mutually independent, squared zero-mean, unit-variance Gaussian random variables is a chi-square random variable with k degrees of freedom.

Chú ý: Tổng của k BNN chuẩn tắc (phương sai=1, kỳ vọng =0) bình phương là một BNN Chi-bình phương với k bậc tự do.

The Gamma Random Variable

BNN Gamma



Other Examples of RVs

Các vị dụ khác về BNN

- **Gamma Random Variable** BNN Gamma
  - Exponential Mũ
  - m-Erlang m-Erlang
  - Chi-Square Chi-bình phương
- **Cauchy Random Variable** BNN Cauchy
- **Rayleigh Random Variable** BNN Rayleigh

THE MARKOV AND CHEBYSHEV INEQUALITIES

Bất đẳng thức Chebyshev và Markov

- Markov inequality Bất đẳng thức Markov

$$P[X \geq a] \leq \frac{E[X]}{a} \qquad \text{for } X \text{ nonnegative.}$$

với X không âm

We obtain Eq. (4.75) as follows:

Chúng ta có được phương tình (4.75) như sau:

$$E[X] = \int_0^a t f_X(t)\, dt + \int_a^\infty t f_X(t)\, dt \geq \int_a^\infty t f_X(t)\, dt$$

$$\geq \int_a^\infty a f_X(t)\, dt = aP[X \geq a].$$

THE MARKOV AND CHEBYSHEV INEQUALITIES

Bất đẳng thức Chebyshev và Markov

- Chebyshev inequality Bất đẳng thức Chebyshev

$$P[|X - m| \geq a] \leq \frac{\sigma^2}{a^2}.$$

Bất đẳng thức Chebyshev là hệ quả của bất đẳng thức Markov. Đặt D2=(...) là bình phương độ lệch khỏi kỳ vòng. Khi đó việc áp bất đẳng thức Markov cho D2 sẽ cho kết quả

The Chebyshev inequality is a consequence of the Markov inequality. Let $D^2 = (X - m)^2$ be the squared deviation from the mean. Then the Markov inequality applied to D^2 gives

$$P[D^2 \geq a^2] \leq \frac{E[(X - m)^2]}{a^2} = \frac{\sigma^2}{a^2}.$$

ENTROPY [OF A RV]

ENTROPY [CỦA 1 BNN] **(Nâng cao, tự tìm hiểu)**

- A measure of the uncertainty in a random experiment. 1 độ đo sự bất định trong 1 thí nghiệm ngẫu nhiên
  - ✓ We are interested in quantifying the uncertainty of the event Ak = {X=k}

Chúng ta quan tâm đến vấn đề lượng hoá sự bất định của một sự kiện Ak = ...

$$I(X = k) = \ln \frac{1}{P[X = k]} = -\ln P[X = k].$$

$$H_X = E[I(X)] = \sum_{k=1}^{K} P[X = k] \ln \frac{1}{P[X = k]}$$

$$= -\sum_{k=1}^{K} P[X = k] \ln P[X = k].$$

Relative Entropy

ENTROPY tương đối **(Nâng cao, tự tìm hiểu)**

Let $\boldsymbol{p} = (p_1, p_2, \ldots, p_K)$, and $\boldsymbol{q} = (q_1, q_2, \ldots, q_K)$ be two pmf's. The **relative entropy** of $\boldsymbol{q}$ with respect to $\boldsymbol{p}$ is defined by

$$H(p;q) = \sum_{k=1}^{K} p_k \ln \frac{1}{q_k} - H_X = \sum_{k=1}^{K} p_k \ln \frac{p_k}{q_k}. \tag{4.107}$$

The relative entropy is nonnegative, and equal to zero if and only if $p_k = q_k$ for all k:

$$H(p;q) \geq 0 \qquad \text{with equality iff} \qquad p_k = q_k \qquad \text{for } k = 1, \ldots, K. \tag{4.108}$$

Entropy for continuous RV

Let X be a continuous random variable that takes on values in the interval $[a, b]$. Suppose that the interval $[a, b]$ is divided into a large number K of subintervals of length Δ. Let $Q(X)$ be the midpoint of the subinterval that contains X. Find the entropy of Q.

Let x_k be the midpoint of the kth subinterval, then $P[Q = x_k] = P[X \text{ is in } k\text{th subinterval}] = P[x_k - \Delta/2 < X < x_k + \Delta/2] \simeq f_X(x_k)\Delta$, and thus

$$H_Q = \sum_{k=1}^{K} P[Q = x_k] \ln P[Q = x_k]$$

$$\simeq -\sum_{k=1}^{K} f_X(x_k)\Delta \ln(f_X(x_k)\Delta)$$

$$= -\ln(\Delta) - \sum_{k=1}^{K} f_X(x_k) \ln(f_X(x_k))\Delta. \tag{4.112}$$

The above equation shows that there is a tradeoff between the entropy of Q and the quantization error $X - Q(X)$. As Δ is decreased the error decreases, but the entropy increases without bound, once again confirming the fact that the entropy of continuous random variables is infinite.

Entropy for continuous RV

- As "Delta" approaches zero, the first expression approaches infinity, but the second expression approaches an integral which may be finite in some cases. The **differential entropy** is defined by this integral:

$$H_X = -\int_{-\infty}^{\infty} f_X(x) \ln f_X(x)\, dx = -E[\ln f_X(X)].$$

Entropy as a Measure of Information

Let X be a discrete random variable with $S_X = \{1, 2, \ldots, K\}$ and pmf $p_k = P[X = k]$. Suppose that the experiment that produces X is performed by John, and that he attempts to communicate the outcome to Mary by answering a series of yes/no questions. We are interested in characterizing the minimum average number of questions required to identify X.

- Read "Example 4.65" on page 207

The Method of Maximum Entropy

- Extra advanced topic. Do it yourself!