

MỘT SỐ ỨNG DỤNG CỦA HÀM SINH XÁC SUẤT

Phạm Thị Cúc¹, Phạm Văn Châu²

TÓM TẮT

Trong bài báo này, chúng tôi giới thiệu một vài ứng dụng của hàm sinh xác suất. Đó là sử dụng hàm sinh xác suất để tính xác suất, kỳ vọng, phương sai, tìm tổng của các biến ngẫu nhiên độc lập, khảo sát quá trình phân nhánh và tính xác suất tuyệt chủng của một họ cá thể.

Từ khóa: Hàm sinh xác suất, biến ngẫu nhiên, phân phối xác suất.

1. ĐẶT VẤN ĐỀ

Hàm sinh được sử dụng rộng rãi trong toán học và đóng vai trò quan trọng trong lý thuyết xác suất [3-5].

Có nhiều loại hàm sinh như : hàm sinh thường, hàm sinh mũ, hàm sinh xác suất, ... Trong bài viết này, chúng tôi sẽ xem xét hàm sinh xác suất đối với các biến ngẫu nhiên rời rạc nhận các giá trị 0,1,2,...

Bản thân tên gọi ‘hàm sinh xác suất’ cũng cho chúng ta một gợi ý về vai trò của nó. Hàm sinh xác suất có thể được sử dụng để ‘sinh ra’ tất cả xác suất của phân phối. Tuy rằng đây không phải là cách hiệu quả được dùng để tính xác suất, sự thực là hàm sinh xác suất đã được chứng minh cho chúng ta biết mọi điều về phân phối xác suất. Bên cạnh đó, hàm sinh xác suất còn được sử dụng để tính kỳ vọng và phương sai.

Trong lý thuyết xác suất, tổng của các biến ngẫu nhiên đóng vai trò đặc biệt quan trọng trong việc nghiên cứu các quá trình ngẫu nhiên, bởi vì nhiều quá trình ngẫu nhiên được tạo nên từ tổng của một dãy các bước lặp, chẳng hạn bài toán về sự thua cuộc của người chơi cờ bạc. Song nhìn chung, việc tìm ra phân phối của tổng bằng cách sử dụng công thức tính xác suất truyền thống là khó khăn. Hàm sinh xác suất được xem là một công cụ hữu dụng để tính tổng và giới hạn của các biến ngẫu nhiên. Ưu điểm đặc biệt của hàm sinh xác suất là cho chúng ta một phương pháp thuận tiện để đặc trưng phân phối tổng $X + Y$ khi X và Y là độc lập. Hàm sinh xác suất chuyển một tổng thành tích và do đó ta có thể tính toán dễ dàng hơn.

Ngoài ra, đối với một số quá trình ngẫu nhiên, hàm sinh xác suất cũng đóng vai trò đặc biệt trong việc chỉ ra khi nào thì quá trình sẽ đạt đến một trạng thái đặc biệt hoặc một trạng thái cho trước.

¹ Giảng viên khoa Khoa học Tự nhiên, Trường Đại học Hồng Đức

² Giáo viên Trường Trung học phổ thông Hậu Lộc 3, huyện Hậu Lộc, tỉnh Thanh Hóa

2. NỘI DUNG

2.1. Hàm sinh xác suất

2.1.1. Định nghĩa và tính chất

Định nghĩa ([5]). Cho X là đại lượng ngẫu nhiên rời rạc nhận giá trị trong tập hợp các số tự nhiên $\{0; 1; 2; 3; \dots\}$. Hàm sinh xác suất của X là hàm số được xác định bởi

$$G_X(x) = p_0 + p_1x + p_2x^2 + \dots = \sum_{k=0}^{\infty} p_k x^k,$$

Trong đó $p_k = P\{X = k\}$ là xác suất để biến ngẫu nhiên X nhận giá trị k .

Chú ý rằng nếu X chỉ nhận một số hữu hạn giá trị thì ta gán cho xác suất bằng 0 đối với những giá trị không thể xảy ra.

Tính chất. Dễ thấy $G_X(x) = E(x^X)$.

Chú ý rằng $G_X(1) = 1$ vì vậy chuỗi hội tụ tuyệt đối với $|x| \leq 1$.

Ta cũng có $G_X(0) = p_0$.

Ta có thể mở rộng định nghĩa hàm sinh xác suất cho các hàm của X . Hàm sinh xác suất của biến ngẫu nhiên rời rạc $Y = f(X)$ là :

$$G_Y(x) = G_{f(X)}(x) = E(x^{f(X)}) = \sum_k P(X = k) x^{f(k)}$$

Nếu Y là hàm đơn giản thì ta có thể biểu diễn $G_Y(x)$ qua $G_X(x)$. Chẳng hạn, nếu $Y = a + bX$ thì:

$$G_Y(x) = E(x^Y) = E(x^{a+bX}) = x^a E[(x^b)^X] = x^a G_X(x^b)$$

2.1.2. Hàm sinh xác suất của một số phân phối thường gặp

i) Phân phối hằng ([1]): Giả sử $X = c$ là biến ngẫu nhiên hằng, nghĩa là $P(X = c) = 1$ và $P(X = k) = 0$ với $k \neq c$. Khi đó hàm sinh xác suất của X là:

$$G_X(x) = E(x^X) = x^c$$

ii) Phân phối Bernoulli ([2]): Giả sử X là biến ngẫu nhiên thỏa mãn: $P(X = 1) = p$, $P(X = 0) = 1 - p = q$, $P(X = k) = 0$ với $k \neq 0, 1$. Khi đó hàm sinh xác suất của X là:

$$G_X(x) = E(x^X) = q + px$$

iii) Phân phối hình học với tham số p ([2]): Giả sử X là biến ngẫu nhiên thỏa mãn $P(X = k) = pq^{k-1}$, $k = 1, 2, \dots$, $q = 1 - p$. Khi đó hàm sinh xác suất của X là:

$$G_X(x) = \frac{px}{1 - qx} \text{ nếu } |x| < \frac{1}{q}$$

iv) Phân phối nhị thức ([2]): Giả sử $X \approx B(n, p)$ là phân phối nhị thức với hai tham số (n, p) , nghĩa là $P(X = k) = C_n^k p^k (1-p)^{n-k}$, $k = 0, 1, 2, \dots, n$. Khi đó hàm sinh xác suất của X là:

$$G_X(x) = (q + px)^n, \quad (q = 1 - p)$$

v) Phân phối Poisson ([2]): Giả sử X là phân phối Poisson với tham số λ , nghĩa là $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$, $k = 0, 1, 2, \dots, n$. Khi đó hàm sinh xác suất của X là:

$$G_X(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k e^{-\lambda} x^k = e^{\lambda(x-1)}$$

2.1.3. Định lý duy nhất

Định lý 1 ([2]). Nếu các biến ngẫu nhiên X và Y có các hàm sinh xác suất lần lượt là $G_X(x)$ và $G_Y(x)$ thì $G_X(x) = G_Y(x)$ với mọi x khi và chỉ khi $P(X = k) = P(Y = k)$ với mọi $k = 0, 1, 2, \dots$. Nghĩa là, $G_X(x) = G_Y(x)$ khi và chỉ khi X và Y có cùng phân phối xác suất.

Chứng minh.

Ta chỉ cần chứng minh điều kiện cần. Do bán kính hội tụ của $G_X(x)$ và $G_Y(x)$ không nhỏ hơn 1 nên chúng có duy nhất một khai triển chuỗi lũy thừa:

$$G_X(x) = \sum_{k=0}^{\infty} P(X = k)x^k, \quad G_Y(x) = \sum_{k=0}^{\infty} P(Y = k)x^k$$

Nếu $G_X(x) = G_Y(x)$ thì hai chuỗi lũy thừa này có các hệ số đồng nhất.

Trong thực hành, nếu ta có thể chỉ ra rằng hai biến ngẫu nhiên có cùng hàm sinh xác suất trên một khoảng nào đó chứa 0 thì hai biến ngẫu nhiên đó có cùng phân phối xác suất. Điều đó có nghĩa là hàm sinh xác suất cho chúng ta biết mọi điều về phân phối xác suất.

2.2. Một số ứng dụng của hàm sinh xác suất

2.2.1. Sử dụng hàm sinh xác suất để tính xác suất

Hàm sinh xác suất có tên gọi như vậy bởi vì chuỗi lũy thừa có thể được khai triển và vi phân của nó sẽ cho chúng ta biết các xác suất riêng biệt. Vì vậy, cho trước một hàm sinh xác suất $G_X(x) = E(x^X)$ ta có thể biết được tất cả các xác suất $p_k = P(X = k)$ như sau:

Từ $G_X(x) = E(x^X) = \sum_{k=0}^{\infty} p_k x^k = p_0 + p_1 x + p_2 x^2 + p_3 x^3 + \dots$ suy ra :

$$p_k = P(X = k) = \frac{1}{k!} G_X^{(k)}(0) = \frac{1}{k!} \frac{d^k}{dx^k} G_X(x) \Big|_{x=0}$$

2.2.2. Tính kỳ vọng và phương sai từ hàm sinh xác suất

Cũng giống như việc tính xác suất, ta có thể sử dụng hàm sinh xác suất để tính kỳ vọng và phương sai của phân phối X .

Định lý 2. Cho X là biến ngẫu nhiên rời rạc có hàm sinh xác suất là $G_X(x)$. Khi đó:

i) $E(X) = G'_X(1)$

ii) $D(X) = G_X^{(2)}(1) - [G_X^{(1)}(1)]^2 + G_X^{(1)}(1)$

Chứng minh.

Ta có: $G'_X(x) = \sum_{k=1}^{\infty} kp_k x^{k-1}$.

Chuỗi này hội tụ với $-1 < x < 1$. Với $x=1$ thì vế phải chính là $\sum kp_k = E(X)$. Nếu kỳ vọng này tồn tại thì đạo hàm $G'_X(x)$ sẽ liên tục trên khoảng đóng $-1 \leq x \leq 1$. Nếu $\sum kp_k$ phân kỳ thì $G'_X(x) \rightarrow \infty$ khi $x \rightarrow 1$. (Trong trường hợp này ta nói X có kỳ vọng vô hạn và viết $G'_X(1) = E(X) = \infty$). Điều này chứng tỏ:

$$E(X) = \sum_{k=1}^{\infty} kp_k = G'_X(1)$$

Tương tự, ta có:

$$E(X(X-1)) = \sum_{k=1}^{\infty} k(k-1)p_k = G_X''(1)$$

Từ đó suy ra phương sai của X là:

$$D(X) = G_X''(1) - [G'_X(1)]^2 + G'_X(1)$$

2.2.3. Tìm tổng của các biến ngẫu nhiên độc lập

2.2.3.1. Tổng của hai biến ngẫu nhiên độc lập

Từ định nghĩa của hàm sinh xác suất $G_X(x) = E(x^X)$ dễ dàng suy ra kết quả sau.

Định lý 3. Giả sử X và Y là hai biến ngẫu nhiên rời rạc độc lập với nhau và có hàm sinh xác suất lần lượt là $G_X(x)$ và $G_Y(x)$. Đặt $Z = X + Y$, khi đó :

$$G_Z(x) = G_{X+Y}(x) = G_X(x)G_Y(x)$$

Hệ quả 4. Nếu X_1, X_2, \dots, X_n là các biến ngẫu nhiên rời rạc độc lập với nhau và có hàm sinh xác suất lần lượt là $G_{X_1}(x), \dots, G_{X_n}(x)$ thì: $G_{X_1+\dots+X_n}(x) = G_{X_1}(x)\dots G_{X_n}(x)$.

Vi dụ 1. Trong một dãy gồm n phép thử độc lập Bernoulli, giả sử $I_i = 1$ nếu phép thử thứ i cho kết quả thành công (với xác suất p) và $I_i = 0$ nếu phép thử thứ i cho kết quả thất bại (với xác suất $q = 1 - p$). Đặt $X = \sum_{i=1}^n I_i$ là số lần thành công trong n phép thử. Tìm phân phối xác suất của X ?

Giải.

Vì các phép thử là độc lập nên các biến ngẫu nhiên I_1, \dots, I_n là độc lập. Vậy

$$G_X(x) = G_{I_1} G_{I_2} \dots G_{I_n}(x)$$

Nhưng $G_{I_i}(x) = q + px$, $i = 1, \dots, n$ nên $G_X(x) = (q + px)^n = \sum_{k=0}^n C_n^k (px)^k q^{n-k}$.

Khi đó, $P(X = k)$ là hệ số của x^k trong $G_X(x)$ và bằng $C_n^k p^k q^{n-k}$, $k = 0, \dots, n$, nghĩa là X là biến ngẫu nhiên có phân phối nhị thức với hai tham số (n, p) .

2.2.3.2. Tổng của một số ngẫu nhiên các biến ngẫu nhiên độc lập

Định lý 5. Giả sử N, X_1, X_2, \dots là các biến ngẫu nhiên rời rạc đếm được. Nếu các biến ngẫu nhiên X_i có cùng phân phối với hàm sinh xác suất là $G_X(x)$ thì $S_N = X_1 + \dots + X_N$ có hàm sinh xác suất là $G_{S_N}(x) = G_N(G_X(x))$. Chú ý rằng ta thừa nhận quy ước $X_1 + \dots + X_N = 0$ với $N = 0$.

Chứng minh.

$$\begin{aligned} \text{Ta có: } G_{S_N}(x) &= E(x^{S_N}) = \sum_{n=0}^{\infty} E(x^{S_N} | N = n) P(N = n) \\ &= \sum_{n=0}^{\infty} E(x^{S_n}) P(N = n) = \sum_{n=0}^{\infty} G_{X_1 + \dots + X_n}(x) P(N = n) \\ &= \sum_{n=0}^{\infty} [G_X(x)]^n P(N = n) = G_N(G_X(x)) \end{aligned}$$

Hệ quả. i) $E(S_N) = E(N).E(X)$

ii) $D(S_N) = E(N)D(X) + D(N)[E(X)]^2$

Ví dụ 2. (Gà Poisson) Một con gà đẻ N quả trứng, với N là biến ngẫu nhiên có phân phối Poisson với tham số λ . Mỗi quả trứng nở ra gà con với xác suất p một cách độc lập với các quả trứng khác. Tìm phân phối xác suất của số gà con được sinh ra Z .

Giải.

Gọi X_i là biến ngẫu nhiên chỉ số gà con được sinh ra từ quả trứng thứ i . Ta có : $Z = X_1 + \dots + X_N$, trong đó X_1, \dots, X_N là các biến ngẫu nhiên có phân phối Bernoulli với tham số p . Khi đó: $G_N(x) = e^{\lambda(x-1)}$, $G_X(x) = q + px$.

Vì vậy $G_Z(x) = G_N(G_X(x)) = e^{\lambda p(x-1)}$, nghĩa là Z là biến ngẫu nhiên có phân phối Poisson với tham số λp .

Ví dụ 3. (Bài toán cờ bạc) Một quý bà trước khi đi du lịch đã nhờ hàng xóm của mình cho cá vàng ở cái ao trong vườn nhà ăn. Mặc dù người hàng xóm sang và cho lũ cá vàng ăn hàng ngày song người này không bao giờ nhìn thấy một con cá vàng nào

trong suốt 3 tuần. Hóa ra tất cả cá vàng đã bị một con cò chén sạch trong khi người này không để ý.

Gọi N là số lần con cò ghé thăm cái ao trong khi người hàng xóm vắng mặt. Giả sử rằng N là một biến ngẫu nhiên có phân phối hình học với tham số $1-\theta$, vì vậy $P(N=n) = (1-\theta)\theta^n$, với $n=0,1,2,\dots$. Khi con cò ghé thăm cái ao, xác suất p để nó bắt được một con cá vàng không phụ thuộc vào các lần ghé thăm khác. (Điều đó có nghĩa là ta đã giả sử có nhiều vô hạn cá vàng để cho cò bắt!). Tìm phân phối xác suất của biến ngẫu nhiên T chỉ tổng số cá vàng bị bắt.

Giải.

Đặt $X_i = 1$ nếu cò bắt được một con cá vàng trong lần ghé thăm và $X_i = 0$ nếu cò không bắt được con cá vàng nào trong lần ghé thăm đó. Khi đó $T = X_1 + X_2 + \dots + X_N$ (tổng của một số ngẫu nhiên các biến ngẫu nhiên).

Vì vậy: $G_T(x) = G_N(G_X(x))$

Bây giờ: $G_X(x) = E(x^X) = x^0.P(X=0) + x^1.P(X=1) = 1-p+px$. Ta cũng có :

$$G_N(y) = \sum_{n=0}^{\infty} y^n P(N=n) = \sum_{n=0}^{\infty} y^n (1-\theta)\theta^n = (1-\theta) \sum_{n=0}^{\infty} (y\theta)^n = \frac{1-\theta}{1-\theta y}, \quad (y < \frac{1}{\theta})$$

$$\text{Suy ra: } G_T(x) = \frac{1-\theta}{1-\theta G_X(x)} = \frac{1-\theta}{1-\theta(1-p+px)} = \frac{1-\theta}{1-\theta+\theta p-\theta px}$$

Tiếp tục biến đổi ta được:

$$G_T(x) = \frac{1-\theta}{(1-\theta+\theta p)-\theta px} = \frac{\frac{1-\theta}{1-\theta+\theta p}}{\frac{(1-\theta+\theta p)-\theta px}{1-\theta+\theta p}} = \frac{1-\theta+\theta p-\theta p}{1-\theta+\theta p} = \frac{1-\frac{\theta p}{1-\theta+\theta p}}{1-\left(\frac{\theta p}{1-\theta+\theta p}\right)x}$$

Đây là hàm sinh xác suất của phân phối hình học với tham số $\left(1-\frac{\theta p}{1-\theta+\theta p}\right)$. Vì vậy

theo tính duy nhất của hàm phân phối xác suất ta kết luận rằng T có phân phối hình học với tham số $\left(1-\frac{\theta p}{1-\theta+\theta p}\right)$.

Nhận xét. Bài toán trên đây cho thấy việc sử dụng hàm sinh xác suất thực sự rất hữu ích trong việc tính tổng của các biến ngẫu nhiên mà nếu sử dụng hàm xác suất thông thường sẽ gặp khó khăn. Thật vậy, để tìm phân phối xác suất của $T = X_1 + X_2 + \dots + X_N$ (là tổng của một số ngẫu nhiên các biến ngẫu nhiên) ta cần tìm biểu thức $P(T=t)$ với t bất kỳ. Để tìm $P(T=t)$ ta phải phân hoạch các giá trị khác nhau của N :

$$P(T=t) = \sum_{n=0}^{\infty} P(T=t|N=n)P(N=n)$$

Sau đó tính tổng theo n . Đối với những bài tổng quát, cả hai bước này nhìn chung sẽ quá khó nếu không có máy tính. Sử dụng hàm sinh xác suất ta không gặp những khó khăn này, và thậm chí nếu $G_T(x)$ không có dạng đơn giản thì ta vẫn có thể biết được mọi điều về phân phối của T .

2.2.4. Sử dụng hàm sinh xác suất để khảo sát quá trình phân nhánh và tìm xác suất tuyệt chủng

Xét một cá thể (giả định) nào đó sống trong đúng một đơn vị thời gian sau đó chết trong quá trình sinh ra một họ các cá thể tương tự. Ta giả sử rằng :

- i) Các kích thước của họ cá thể là các biến ngẫu nhiên độc lập nhận các giá trị $0, 1, 2, \dots$;
- ii) Các kích thước của họ cá thể là những biến ngẫu nhiên có cùng phân phối xác suất, số ‘con’ trong một họ ký hiệu là C có phân phối : $P(C = k) = p_k, k = 0, 1, 2, \dots$

Sự phát triển của ‘dân số’ khi thời gian tiếp diễn gọi là *quá trình phân nhánh*. Giả sử X_n là số cá thể được sinh ra ở thời điểm n (nghĩa là kích thước của thế hệ thứ n). Sự phát triển của dân số được mô tả bởi một dãy các biến ngẫu nhiên X_0, X_1, X_2, \dots . Giả sử rằng $X_0 = 1$, nghĩa là ta bắt đầu với đúng một cá thể.

Ta có thể sử dụng hàm sinh xác suất để khảo sát quá trình này như sau. Gọi $G(x)$ là hàm sinh xác suất của C , theo định nghĩa hàm sinh xác suất thì : $G(x) = \sum_{k=0}^{\infty} P(C = k)x^k$

và gọi $G_n(x)$ là hàm sinh xác suất của X_n thì: $G_n(x) = \sum_{i=0}^{\infty} P(X_n = i)x^i$

Bây giờ, giả sử $G_0(x) = x$, vì $P(X_0 = 1) = 1; P(X_0 = i) = 0$ với $i \neq 1$ và $G_1(x) = G(x)$.

Ta cũng có : $X_n = C_1 + C_2 + \dots + C_{n-1}$

Trong đó: C_j là kích thước của họ được sinh ra bởi thành viên thứ j của thế hệ thứ $(n-1)$. Như vậy, do X_n là tổng của một số ngẫu nhiên các biến ngẫu nhiên độc lập và có cùng phân phối xác suất nên ta có: $G_n(x) = G_{n-1}(G(x))$ với $n = 2, 3, \dots$ và điều này cũng đúng với $n = 1$.

Lặp lại công thức này, ta được :

$$G_n(x) = G_{n-1}(G(x)) = G_{n-2}(G(G(x))) = \dots = G_1(G(G(\dots(x)\dots))) = G(G(G(\dots(x)\dots)))$$

với $n = 0, 1, 2, \dots$, nghĩa là G_n là bước lặp thứ n của G .

Bây giờ, vì $G(1) = 1$ nên ta có : $E(X_n) = G'_n(1) = G'_{n-1}(G(1))G'(1) = G'_{n-1}(1)G'(1)$, nghĩa là $E(X_n) = E(X_{n-1})\mu$, trong đó $\mu = E(C)$ là kích thước trung bình của họ cá thể. Do đó :

$$E(X_n) = \mu E(X_{n-1}) = \mu^2 E(X_{n-2}) = \dots = \mu^n E(X_0) = \mu^n$$

$$\text{Vì vậy : } E(X_n) \rightarrow \begin{cases} 0, & \mu < 1 \\ 1, & \mu = 1 \\ \infty, & \mu > 1 \end{cases}$$

Thực ra ta có thể chỉ ra được toàn bộ phân phối xác suất của X_n bằng cách khai triển hàm sinh dưới dạng chuỗi lũy thừa của x , song kết quả này khá phức tạp. Tuy nhiên, trường

$$\text{hợp đặc biệt : } P(X_n = 0) = G_n(0) = p \frac{q^n - p^n}{q^{n+1} - p^{n+1}} = \frac{\mu^n - 1}{\mu^{n+1} - 1}$$

Trong đó $\mu = q/p$ là kích thước trung bình của họ. Từ đó suy ra sự tuyệt chủng là chắc chắn nếu $\mu < 1$ và ít chắc chắn hơn nếu $\mu > 1$.

Dưới đây, ta sẽ chỉ ra rằng kết luận trên đây về sự tuyệt chủng là đúng đối với tất cả các phân phối của kích thước họ cá thể.

Xác suất để quá trình tuyệt chủng sau n thế hệ là $e_n = P(X_n = 0)$. Ta có $e_n \leq 1$ và $e_n \leq e_{n+1}$ (vì $X_n = 0$ kéo theo $X_{n+1} = 0$), nghĩa là $\{e_n\}$ là một dãy đơn điệu bị chặn. Vì vậy $e = \lim_{n \rightarrow \infty} e_n$ tồn tại và được gọi là *xác suất tuyệt chủng*.

Định lý 6. e là nghiệm không âm nhỏ nhất của phương trình $x = G(x)$.

Chứng minh.

Chú ý rằng: $e_n = P(X_n = 0) = G_n(0)$. Ta có:

$$G_n(x) = G_{n-1}(G(x)) = \dots = G(G \dots (x) \dots) = G(G_{n-1}(x))$$

Cho $x = 0$ thì $e_n = G_n(0) = G(e_{n-1})$, $n = 1, 2, \dots$, với điều kiện biên $e_0 = 0$. Lấy giới hạn khi $n \rightarrow \infty$ ta được $e = G(e)$.

Bây giờ giả sử η là nghiệm không âm bất kỳ của phương trình $x = G(x)$. Do $G(x)$ là không giảm trên đoạn $[0; 1]$ (vì nó có các hệ số không âm) nên

$$e_1 = G(e_0) = G(0) \leq G(\eta) = \eta$$

$$e_2 = G(e_1) \leq G(\eta) = \eta$$

và bằng qui nạp ta có $e_n \leq \eta$ với $n = 1, 2, \dots$

Vì vậy : $e = \lim_{n \rightarrow \infty} e_n \leq \eta$. Suy ra e là nghiệm không âm nhỏ nhất.

Định lý 7. $e = 1$ khi và chỉ khi $\mu \leq 1$.

Chứng minh.

Ta có thể giả sử rằng $p_0 > 0$ (vì nếu ngược lại thì $e = 0$ và $\mu > 1$). Bây giờ, trên đoạn $[0; 1]$ thì $G(x)$ là :

liên tục (vì bán kính hội tụ ≥ 1)

không giảm (vì $G'(x) = \sum_k k p_k x^{k-1} \geq 0$)

lồi (vì $G''(x) = \sum_k k(k-1) p_k x^{k-2} \geq 0$).

Từ đó suy ra rằng trong đoạn $[0;1]$ thì đường thẳng $y = x$ có một (nếu $G'(1) \leq 1$) hoặc hai giao điểm (nếu $G'(1) > 1$) với đường cong $y = G(x)$.

Vì $G'(1) = \mu$ nên suy ra $e = 1$ khi và chỉ khi $\mu \leq 1$.

Ví dụ 4. Tìm xác suất tuyệt chủng khi C có phân phối kiểu hình học

$$p_k = pq^k, \quad k = 0, 1, 2, \dots; \quad 0 < p = 1 - q < 1.$$

Giải.

Hàm sinh xác suất của C là: $G(x) = \frac{p}{1 - qx}$, $|x| < \frac{1}{q}$

và e là nghiệm không âm nhỏ nhất của phương trình: $G(x) = \frac{p}{1 - qx} = x$

Các nghiệm của phương trình này là $\frac{1 \pm (2p - 1)}{2(1 - p)}$.

Vi vậy:

nếu $p < \frac{1}{2}$ (nghĩa là $\mu = \frac{q}{p} > 1$) thì $e = \frac{p}{q} (= \mu^{-1})$,

nếu $p \geq \frac{1}{2}$ (nghĩa là $\mu \leq 1$) thì $e = 1$.

3. KẾT LUẬN

Trên đây, sau khi giới thiệu về hàm sinh xác suất cùng với một số tính chất và ví dụ điển hình, chúng tôi đã trình bày một số ứng dụng cơ bản của hàm sinh trong lý thuyết xác suất. Ngoài các ứng dụng này, hàm sinh xác suất cũng có thể được sử dụng để tính mô-men, giải hệ thức đệ quy hoặc xem xét khi nào thì một quá trình đạt đến một trạng thái đặc biệt nào đó.

TÀI LIỆU THAM KHẢO

- [1] Đào Hữu Hồ (2006), *Xác suất thống kê*, Nxb. Đại học Quốc gia Hà Nội, Hà Nội.
- [2] Đỗ Đức Thái, Nguyễn Tiến Dũng (2010), *Nhập môn hiện đại Xác suất và Thống kê*, Nxb. Đại học Sư phạm, Hà Nội.
- [3] Feller W. (1968), *An introduction to probability theory and its applications*, Vol.1, 3rd Edition, John Wiley, New York.
- [4] Malliavin, P. (1995), *Integration and Probability*, Springer Verlag.
- [5] Manuel, L. Ésquivel (2009), *Some applications of probability generating function based methods to statistical estimation*, *Discussiones Mathematicae Probability and Statistics*, 29, 131-153.

SOME APPLICATIONS OF PROBABILITY GENERATING FUNCTIONS

Pham Thi Cuc, Pham Van Chau

ABSTRACT

In this paper, we introduce some applications of probability generating functions: using probability generating functions to calculate probability, expectation, variance, sum of independent random variables, investigating branching processes and calculating the extinction probability of a family of organisms.

Keywords: *Probability generating function, random variable, probability distribution.*