Chương 2
BIẾN NGẪU NHIÊN VÀ LUẬT PHÂN PHỐI XÁC SUẤT

BỘ MÔN TOÁN ỨNG DỤNG[1]

VIỆN TOÁN ỨNG DỤNG VÀ TIN HỌC
ĐẠI HỌC BÁCH KHOA HÀ NỘI

SAMI.HUST – 2023

[1] Phòng BIS.201-D3.5

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối đều rời rạc

Định nghĩa 11

Biến ngẫu nhiên rời rạc X được gọi là có phân phối đều với tham số n nếu mỗi giá trị trong n giá trị có thể có của nó, $x_1, x_2, \ldots, x_n$, có xác suất bằng nhau,

$$P(X = x_i) = \frac{1}{n}, \quad \forall i = 1, 2, \ldots, n.$$

Ví dụ

Ví dụ 29

Chữ số đầu tiên của số sê-ri của một sản phẩm là một trong các chữ số từ 0 đến 9. Lấy ngẫu nhiên một sản phẩm từ lô hàng chứa nhiều sản phẩm và gọi X là chữ số đầu tiên của số sê-ri của sản phẩm được lấy ra. Khi đó, X có phân phối đều với tham số $n = 10$.

- Hàm xác suất của X là

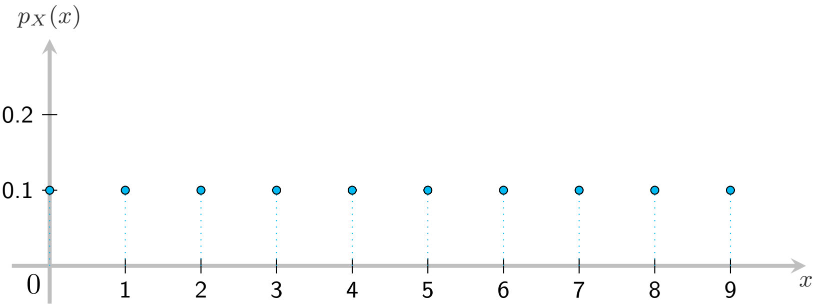
$$p_X(x) = \begin{cases} 0,1, & x = 0, 1, \ldots, 9, \\ 0, & \text{nếu trái lại.} \end{cases}$$

- Bảng phân phối xác suất của X là

X	0	1	2	3	4	5	6	7	8	9
$P(X = x_i)$	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1

Đồ thị phân phối xác suất của X được cho trong Hình 9.

Ví dụ



Hình 9: Hàm xác suất của biến ngẫu nhiên có phân phối đều trong Ví dụ 29

Kỳ vọng và phương sai

Định lý 5

Giả sử X là một biến ngẫu nhiên có phân phối đều rời rạc nhận các giá trị nguyên liên tiếp $a, a+1, a+2, \ldots, b$, $(a < b)$. Khi đó, kỳ vọng và phương sai của X là

$$E(X) = \frac{a+b}{2} \quad \text{và} \quad V(X) = \frac{(b-a+1)^2 - 1}{12}. \tag{29}$$

Ví dụ 30

Hệ thống liên lạc bằng giọng nói của một doanh nghiệp có 48 đường truyền. Quan sát hệ thống tại một thời điểm cụ thể và gọi X là "số đường truyền đang được sử dụng tại thời điểm quan sát". Giả sử X là một biến ngẫu nhiên có phân phối đều, tìm $E(X)$ và $V(X)$.

Kỳ vọng và phương sai

Giải. X nhận các giá trị từ 0 đến 48. Khi đó, theo (29),

$$E(X) = \frac{48+0}{2} = 24; \quad V(X) = \frac{(48-0+1)^2-1}{12} = 200.$$

Suy ra $\sigma_X = \sqrt{V(X)} = 14,1421$. Ta thấy số đường truyền trung bình được sử dụng là 24 nhưng độ phân tán, được đo bằng σ_X, là lớn. Vì vậy, tại nhiều thời điểm, sẽ có nhiều hơn hoặc ít hơn 24 đường truyền được sử dụng.

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối Bernoulli

Định nghĩa 12

Biến ngẫu nhiên rời rạc X được gọi là có phân phối Bernoulli với tham số p, ký hiệu là $X \sim \mathcal{B}(p)$, nếu X chỉ nhận hai giá trị 0 và 1 và hàm xác suất của X là

$$p_X(x) = \begin{cases} 1-p, & x = 0, \\ p, & x = 1, \end{cases} \tag{30}$$

với $p \in (0; 1)$.

✍ Xét một phép thử Bernoulli với xác suất xuất hiện sự kiện A trong phép thử này là $p(A) = p \in (0; 1)$. Gọi X là số lần xuất hiện A trong phép thử Bernoulli này, thì X là biến ngẫu nhiên rời rạc có phân phối Bernoulli tham số p.

Kỳ vọng và phương sai

Định lý 6

Nếu X là biến ngẫu nhiên có phân phối Bernoulli tham số p thì

$$E(X) = p \quad \text{và} \quad V(X) = p(1-p). \tag{31}$$

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối nhị thức

Định nghĩa 13

Biến ngẫu nhiên rời rạc X được gọi là có phân phối nhị thức với tham số n và p, ký hiệu là $X \sim \mathcal{B}(n;p)$, nếu X nhận hữu hạn giá trị $0, 1, \ldots, n$ với hàm xác suất

$$p_X(x) = C_n^x p^x (1-p)^{n-x}, \; x = 0, 1, \ldots, n, \tag{32}$$

với $p \in (0;1)$.

Định nghĩa phân phối nhị thức

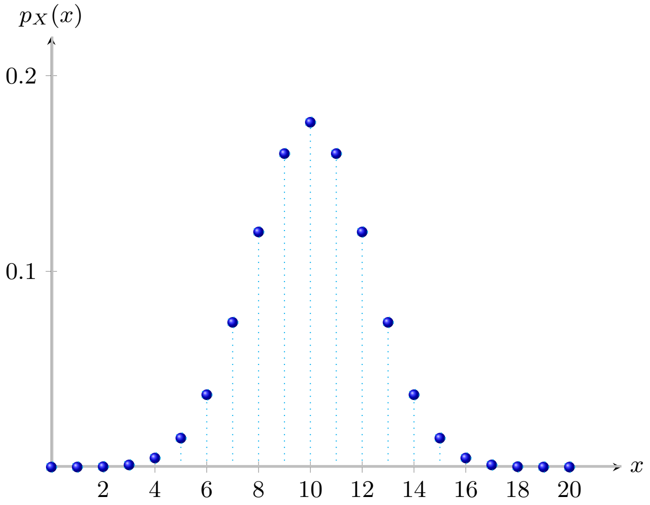
✍

- Phân phối nhị thức xuất phát từ tên thực tế của khai triển nhị thức $(p+q)^n$ có $n+1$ số hạng:

$$(p+q)^n = C_n^0 p^0 q^n + C_n^1 p^1 q^{n-1} + \cdots + C_n^n p^n q^0.$$

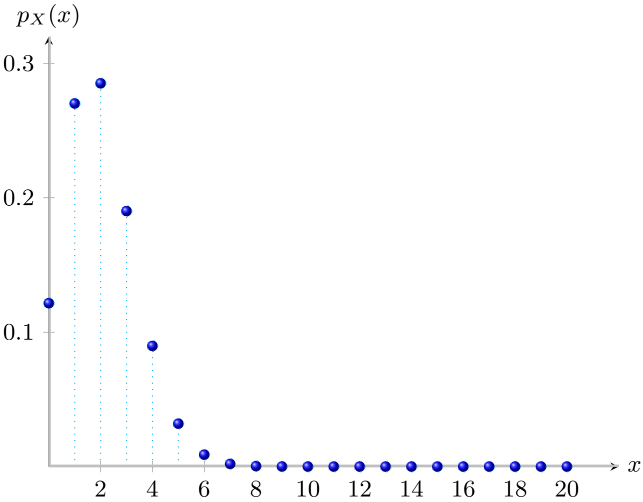
- Nếu $p+q=1$, tức là $q=1-p$, thì $\sum_{k=0}^{n} C_n^k p^k q^{n-k} = 1$, đây là điều kiện cần thiết của phân phối nhị thức.
- Trong n phép thử Bernoulli, nếu X là biến ngẫu nhiên đếm số lần xuất hiện sự kiện A với $P(A) = p \in (0;1)$ không đổi, thì $X \sim \mathcal{B}(n;p)$.

Định nghĩa phân phối nhị thức



Hình 10: Phân phối nhị thức với $n = 20$ và $p = 0,5$

Định nghĩa phân phối nhị thức



Hình 11: Phân phối nhị thức với $n = 20$ và $p = 0,1$

Ví dụ

Ví dụ 31

Một lô hàng có số lượng lớn sản phẩm với tỷ lệ sản phẩm A trong đó là 50%. Từ lô hàng, chọn ngẫu nhiên 20 sản phẩm để kiểm tra. Gọi X là số sản phẩm A có trong 20 sản phẩm được lấy ra.

(a) X có phân phối gì?

(b) Tính xác suất có đúng 5 sản phẩm A trong 20 sản phẩm được lấy ra.

Giải. Có thể xem việc kiểm tra chất lượng mỗi sản phẩm là thực hiện một phép thử Bernoulli với xác suất xuất hiện A trong mỗi phép thử là 0,5 không đổi. Kiểm tra 20 sản phẩm là thực hiện 20 phép thử Bernoulli.

(a) X là biến ngẫu nhiên đếm số lần xuất hiện A trong 20 phép thử Bernoulli, nên $X \sim \mathcal{B}(n;p)$ với $n = 20$ và $p = 0,5$ (Hình 10).

(b) Áp dụng công thức Bernoulli,

$$P(X = 5) = C_{20}^{5} \times (0,5)^{5} \times (1 - 0,5)^{15} \approx 0,0148.$$

Kỳ vọng và phương sai

Định lý 7

Nếu X là biến ngẫu nhiên có phân phối nhị thức với tham số n và p thì

(a) Kỳ vọng và phương sai của X là

$$E(X) = np \quad \text{và} \quad V(X) = np(1-p). \tag{33}$$

(b) $\mathrm{mod}(X)$ là giá trị của X thỏa mãn

$$(n+1)p - 1 \leq \mathrm{mod}(X) \leq (n+1)p. \tag{34}$$

Kỳ vọng và phương sai

Ví dụ 32

- Biến ngẫu nhiên X trong Ví dụ 5 là biến ngẫu nhiên có phân phối nhị thức với tham số $n = 4$ và $p = 0,1$.
- Áp dụng (33),

$$E(X) = np = 4 \times 0,1 = 0,4 \quad \text{và} \quad V(X) = np(1-p) = 4 \times 0,1 \times 0,9 = 0,36.$$

- Những kết quả này trùng với những kết quả thu được từ việc tính toán trực tiếp trong Ví dụ 18(a) và 23(a).

Ví dụ 33

- Trong Ví dụ 31, X là biến ngẫu nhiên có phân phối nhị thức với $n = 20$ và $p = 0,5$.
- Vì $(n+1)p - 1 = 9,5$ và $(n+1)p = 10,5$ nên $\mathrm{mod}(X) = 10$.
- Số sản phẩm A có nhiều khả năng nhất là 10 trong số 20 sản phẩm được lấy ra.

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối Poisson

Định nghĩa 14

Biến ngẫu nhiên rời rạc X được gọi là có phân phối Poisson với tham số $\lambda > 0$, ký hiệu là $X \sim \mathcal{P}(\lambda)$, nếu hàm xác suất của X có dạng

$$p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \ x = 0, 1, 2, \ldots \tag{35}$$

với $e = 2,71828\ldots$

Định nghĩa phân phối Poisson

✍

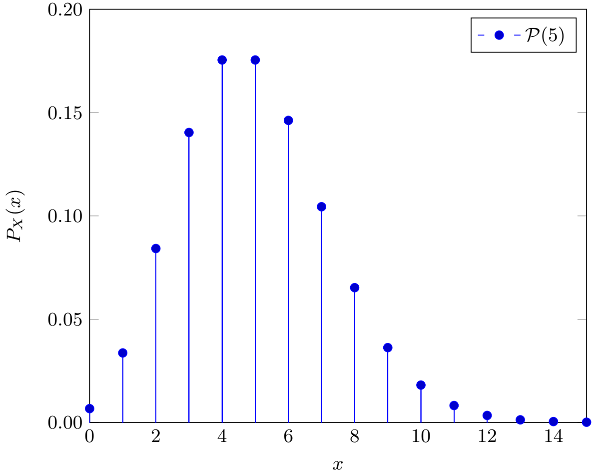
- Tổng các xác suất trong (35) bằng một vì

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1,$$

  ở đây ta sử dụng khai triển Taylor của hàm $e^{\lambda} = \sum_{x=0}^{\infty} \lambda^x/x!$.

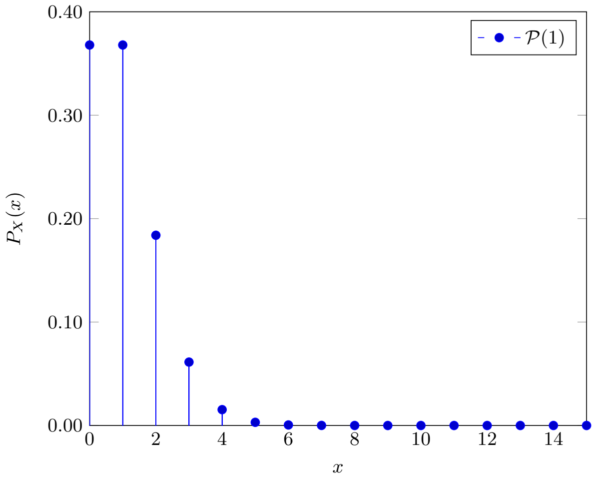
- Trong thực tế với một số giả thiết thích hợp thì các biến ngẫu nhiên là các quá trình đếm sau: số cuộc gọi đến một tổng đài; số khách hàng đến một điểm phục vụ; số xe cộ qua một ngã tư; số tai nạn (xe cộ); số các sự cố xảy ra ở một địa điểm ... trong một khoảng thời gian xác định nào đó sẽ có phân phối Poisson với tham số λ, trong đó λ là giá trị trung bình diễn ra trong khoảng thời gian này.

Định nghĩa phân phối Poisson



Hình 12: Phân phối Poisson với $\lambda = 5$

Định nghĩa phân phối Poisson



Hình 13: Phân phối Poisson với $\lambda = 1$

Ví dụ

Ví dụ 34

Ở một tổng đài bưu điện, các cuộc điện thoại gọi đến xuất hiện ngẫu nhiên, độc lập với nhau với tốc độ trung bình 2 cuộc gọi trong một phút. Tìm xác suất để

(a) Có đúng 6 cuộc điện thoại trong vòng 2,5 phút.

(b) Không có cuộc điện thoại nào trong khoảng thời gian 30 giây.

(c) Có ít nhất 1 cuộc điện thoại trong khoảng thời gian 10 giây.

Ví dụ

Giải.

(a) Gọi X là "số cuộc điện thoại xuất hiện trong vòng 2,5 phút". Khi đó, $X \sim \mathcal{P}(\lambda)$ với $\lambda = 5$. Phân phối xác suất của X xác định bởi (35) cho trong Hình 12 và

$$P(X = 6) = e^{-5}\frac{5^6}{6!} \approx 0,1462.$$

(b) Gọi Y là "số cuộc điện thoại xuất hiện trong vòng 30 giây". Khi đó, $Y \sim \mathcal{P}(\lambda)$ với $\lambda = 1$. Phân phối xác suất của Y xác định bởi (35) cho trong Hình 13 và

$$P(Y = 0) = e^{-1}\frac{\lambda^0}{0!} \approx 0,3679.$$

(c) Gọi Z là "số cuộc điện thoại xuất hiện trong vòng 10 giây". Khi đó, $Z \sim \mathcal{P}(\lambda)$ với $\lambda = 1/3$ và

$$P(Z \geq 1) = 1 - P(Z = 0) = 1 - e^{-1/3} \approx 0,2835.$$

Kỳ vọng và phương sai

Định lý 8

Nếu X là biến ngẫu nhiên có phân phối Poisson tham số λ thì

$$E(X) = \lambda \quad \text{và} \quad V(X) = \lambda. \tag{36}$$

Xấp xỉ phân phối nhị thức bằng phân phối Poisson

✍ Cho X là biến ngẫu nhiên có phân phối nhị thức với tham số n và p. Nếu $n \to \infty$, $p \to 0$ và $np \to \lambda$ (λ là một hằng số) thì

$$\mathcal{B}(n;p) \to \mathcal{P}(\lambda).$$

✍ Trong thực tế, nếu n đủ lớn và p đủ nhỏ thỏa mãn $np < 7$ thì ta có thể xấp xỉ phân phối nhị thức $\mathcal{B}(n;p)$ bằng phân phối Poisson $\mathcal{P}(\lambda)$ với $\lambda = np$ và

$$P_n(k) = C_n^k p^k (1-p)^{n-k} \approx \frac{(np)^k}{k!} e^{-np}. \tag{37}$$

Xấp xỉ phân phối nhị thức bằng phân phối Poisson

Ví dụ 35

Giả sử một công ty bảo hiểm nhân thọ bảo hiểm cho cuộc sống của 5000 người đàn ông ở độ tuổi 42. Nghiên cứu của các chuyên gia tính toán cho thấy xác suất để một người đàn ông 42 tuổi sẽ chết trong một năm (xác định) là 0,001. Hãy tìm xác suất mà công ty sẽ phải trả bảo hiểm cho 4 người trong một năm (xác định).

Giải.

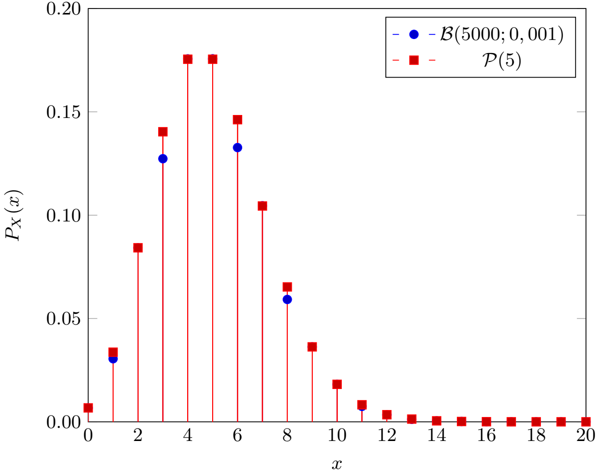
- Gọi X là số người chết trong một năm (xác định), $X \sim \mathcal{B}(n; p)$ với $n = 5000$ và $p = 0,001$. Khi đó, áp dụng công thức Bernoulli,

$$P(X = 4) = P_{5000}(4) = C_{5000}^{4}(0,001)^4(1-0,001)^{5000-4} = \frac{5000!}{4!4996!}(0,001)^4(0,999)^{4996} = 0,17552002723.$$

- Vì $n = 5000$ đủ lớn và $\lambda = np = (5000)(0,001) = 5 < 7$ nên xác suất trên có thể được xấp xỉ bằng công thức (37) và

$$P(X = 4) \approx \frac{5^4}{4!}e^{-5} = \frac{(625)(0,006738)}{24} = 0,175.$$

Xấp xỉ phân phối nhị thức bằng phân phối Poisson



Hình 14: Xấp xỉ phân phối nhị thức $\mathcal{B}(5000; 0,001)$ bằng phân phối Poisson $\mathcal{P}(5)$

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối đều liên tục

Định nghĩa 15

Biến ngẫu nhiên liên tục X được gọi là có phân phối đều trên $[a; b]$ $(a < b)$, ký hiệu là $X \sim \mathcal{U}([a; b])$, nếu X có hàm mật độ xác suất là

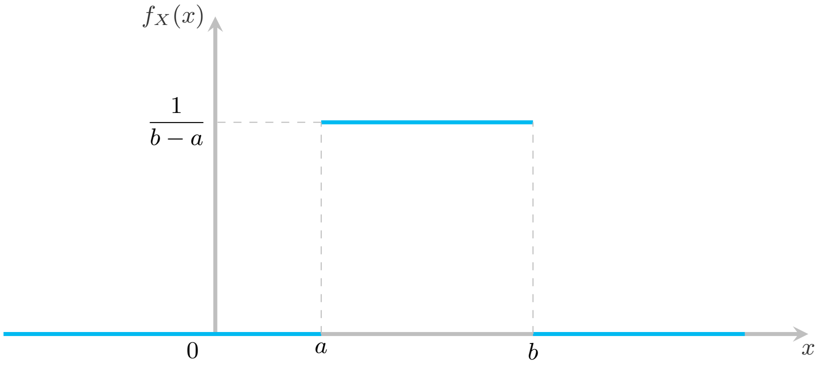
$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a; b], \\ 0, & x \notin [a; b]. \end{cases} \tag{38}$$

Nhận xét 1

Từ Định nghĩa 15 và tính chất của hàm mật độ xác suất suy ra nếu $X \sim \mathcal{U}([a; b])$ và nếu $\alpha, \beta \in [a, b]$, thì

$$P(\alpha < X < \beta) = \frac{\beta - \alpha}{b - a}. \tag{39}$$

Định nghĩa phân phối đều liên tục



Hình 15: Hàm mật độ xác suất của biến ngẫu nhiên có phân phối đều trên đoạn $[a; b]$

Ví dụ

Ví dụ 36

Tại một trạm xe buýt, chuyến xe đầu tiên khởi hành lúc 7 giờ sáng và cứ sau 15 phút lại có một xe xuất bến. Giả sử một hành khách đến trạm ngẫu nhiên trong khoảng thời gian từ 7 giờ đến 7 giờ 30. Tìm xác suất để hành khách này chờ

(a) Ít hơn 5 phút.

(b) Ít nhất 12 phút.

Ví dụ

Giải. Quy 7 giờ về 0 và 7 giờ 30 về 30 (phút). Gọi X là thời điểm khách đến trạm từ 7 giờ đến 7 giờ 30, $X \sim \mathcal{U}[0;30]$.

(a) Hành khách chờ ít hơn 5 phút nếu đến trạm giữa 7 giờ 10 và 7 giờ 15 hoặc giữa 7 giờ 25 và 7 giờ 30. Do đó xác suất cần tìm là:

$$P(10 < X \leq 15) + P(25 < X \leq 30) = \frac{15-10}{30-0} + \frac{30-25}{30-0} = \frac{1}{3}.$$

(b) Hành khách chờ ít nhất 12 phút nếu đến trạm giữa 7 giờ và 7 giờ 03 hoặc giữa 7 giờ 15 và 7 giờ 18. Xác suất cần tìm là:

$$P(0 < X \leq 3) + P(15 < X \leq 18) = \frac{3-0}{30-0} + \frac{18-15}{30-0} = 0,2.$$

Kỳ vọng và phương sai

Định lý 9

Cho X là biến ngẫu nhiên có phân phối đều trên $[a;b]$. Khi đó,

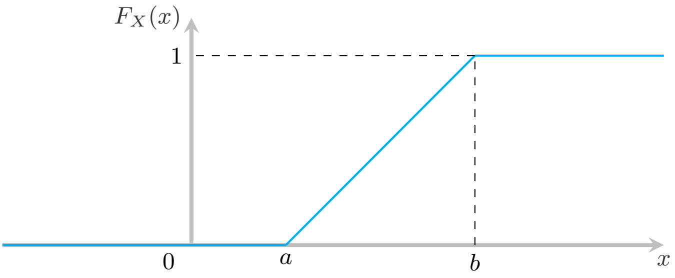
(a) Hàm phân phối xác suất của X là

$$F_X(x) = \begin{cases} 0, & x \le a, \\ \dfrac{x-a}{b-a}, & a < x \le b, \\ 1, & x > b. \end{cases} \tag{40}$$

(b) Kỳ vọng $E(X) = \dfrac{a+b}{2}$.

(c) Phương sai $V(X) = \dfrac{(b-a)^2}{12}$ và độ lệch chuẩn $\sigma_X = \sqrt{\dfrac{(b-a)^2}{12}}$.

b. Kỳ vọng và phương sai



Hình 16: Hàm phân phối xác suất của biến ngẫu nhiên có phân phối đều trên đoạn $[a; b]$

Kỳ vọng và phương sai

Ví dụ 37

Biến ngẫu nhiên X trong Ví dụ 15 có phân phối đều trên đoạn $[0; 20]$, do đó

$$E(X) = \frac{0+20}{2} = 10 \text{ (mA)} \quad \text{và} \quad V(X) = \frac{(20-0)^2}{12} = 33,3333 \text{ (mA}^2\text{)}.$$

Kết quả này trùng với kết quả tính trực tiếp trong Ví dụ 18(c) và 23(b).

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

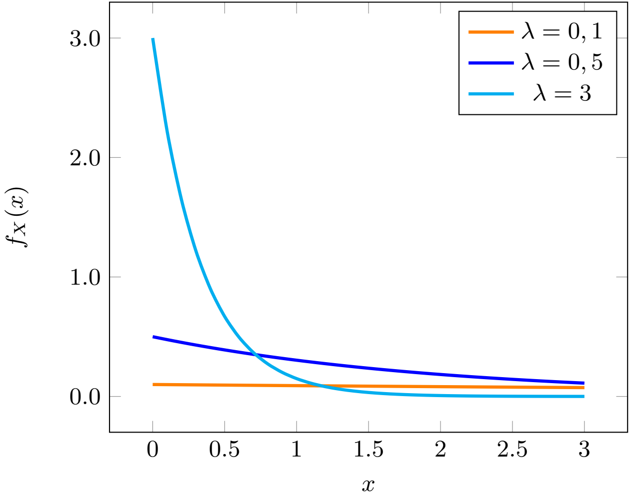
Định nghĩa phân phối mũ

Định nghĩa 16

Biến ngẫu nhiên X được gọi là có phân phối mũ với tham số $\lambda > 0$, ký hiệu là $X \sim \text{Exp}(\lambda)$, nếu hàm mật độ xác suất của nó có dạng

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{41}$$

Định nghĩa phân phối mũ



Hình 17: Hàm mật độ của biến ngẫu nhiên tuân có phân phối mũ với tham số khác nhau

Kỳ vọng và phương sai

Định lý 10

Cho X là biến ngẫu nhiên có phân phối mũ với tham số $\lambda > 0$. Khi đó,

(a) $F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$

(b) $E(X) = 1/\lambda$.

(c) $V(X) = 1/\lambda^2$.

✍ Từ Định lý 10(a) suy ra $P(X > x) = 1 - P(X \leq x) = 1 - P(X < x) = 1 - F_X(x)$. Do đó, nếu $X \sim \text{Exp}(\lambda)$ thì

$$P(X > x) = e^{-\lambda x}, \quad x \geq 0. \tag{42}$$

Ví dụ

Ví dụ 38

Biến ngẫu nhiên T có phân phối mũ với tham số $\lambda > 0$ có hàm phân phối xác suất là

$$F_T(t) = \begin{cases} 1 - e^{-t/3}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

(a) Tìm hàm mật độ xác suất của T.

(b) Tính $P(2 \leq T \leq 4)$.

(c) Tính kỳ vọng, phương sai, độ lệch chuẩn của T.

Ví dụ

Giải.

(a) Ta có

$$f_T(t) = \frac{dF_T(t)}{dt} = \begin{cases} \frac{1}{3}e^{-t/3}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Theo Định nghĩa 16, $T \sim \text{Exp}(\lambda)$ với $\lambda = 1/3$.

(b) Sử dụng tính chất của hàm phân phối,

$$P(2 \leq T \leq 4) = F_T(4) - F_T(2) = e^{-2/3} - e^{-4/3} \approx 0,250.$$

(c) Sử dụng Định lý 10(b),(c) ta nhận được

$$E(T) = \frac{1}{1/3} = 3, \quad V(T) = \frac{1}{(1/3)^2} = 9 \quad \text{và} \quad \sigma_T = \sqrt{9} = 3.$$

Ví dụ

Ví dụ 39

Giả sử tuổi thọ (năm) của một mạch điện tử trong máy tính là một biến ngẫu nhiên có phân phối mũ với kỳ vọng là 6,25 năm. Thời gian bảo hành của mạch điện tử này là 5 năm. Hỏi có bao nhiêu phần trăm mạch điện tử bán ra phải thay thế trong thời gian bảo hành.

Giải. Gọi X là tuổi thọ của mạch điện tử trong máy tính, khi đó $X \sim \text{Exp}(\lambda)$ với $\lambda = \dfrac{1}{E(X)} = \dfrac{1}{6,25} = 0,16$. Do đó,

$$P(X \leq 5) = 1 - e^{-5\lambda} = 1 - e^{-0,8} \simeq 0,5506.$$

Vậy có khoảng 55,06% mạch điện tử bán ra phải thay thế trong thời gian bảo hành.

Tính chất không nhớ

Ví dụ 40

Một động cơ có thời gian làm việc (giờ) liên tục cho đến khi có hỏng hóc đầu tiên được giả thiết là biến ngẫu nhiên có phân phối mũ với tham số $\lambda = 0,001$.

(a) Tính xác suất để động cơ này làm việc liên tục không hỏng hóc được ít nhất 900 giờ.

(b) Được biết động cơ này đã làm việc liên tục không hỏng hóc được ít nhất 300 giờ, xác suất để động cơ còn làm việc liên tục không hỏng hóc được ít nhất 1200 giờ là bao nhiêu?

Giải. Gọi X là thời gian làm việc của động cơ, $X \sim \text{Exp}(\lambda)$ với $\lambda = 0,001$.

(a) Sử dụng (42), $P(X > 900) = e^{-0,001 \times 900} = e^{-0,9} \approx 0,4066$.

(b) Áp dụng công thức xác suất điều kiện và (42),

$$P\big((X > 1200)|(X > 300)\big) = \frac{P(X > 1200)}{P(X > 300)} = e^{-0,001 \times (1200 - 300)} \approx 0,4066.$$

Tính chất không nhớ

✍ Ví dụ 40 minh họa "tính chất không nhớ" (Lack of Memory Property) của một biến ngẫu nhiên có phân phối mũ.

Tính chất 6

Với một biến ngẫu nhiên X có phối mũ tham số $\lambda > 0$,

$$P(X > s + t | X > t) = P(X > s). \tag{43}$$

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối chuẩn

Định nghĩa 17

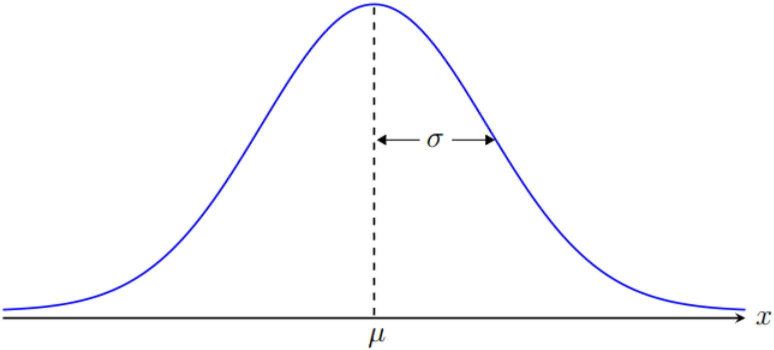
Biến ngẫu nhiên liên tục X được gọi là có phân phối chuẩn với tham số μ và σ^2, ký hiệu là $X \sim \mathcal{N}(\mu; \sigma^2)$, nếu hàm mật độ xác suất của X có dạng

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}} \tag{44}$$

ở đây, tham số μ là số thực tùy ý còn $\sigma > 0$.

✍Đồ thị của hàm mật độ xác suất $f_X(x)$ của biến ngẫu nhiên có phân phối chuẩn, được gọi là đường cong chuẩn, có dạng hình chuông (Hình 18).

Định nghĩa phân phối chuẩn



Hình 18: Đường cong chuẩn

Định nghĩa phân phối chuẩn

✍

- Phân phối chuẩn được Gauss tìm ra năm 1809 nên nó còn được gọi là phân phối Gauss.
- Trong thực tế, với một số điều kiện thích hợp, nhiều biến ngẫu nhiên có phân phối chuẩn hoặc tiệm cận chuẩn. Chẳng hạn, trọng lượng, chiều cao của một nhóm người nào đó; điểm thi của thí sinh; năng suất cây trồng . . . .

Kỳ vọng và phương sai

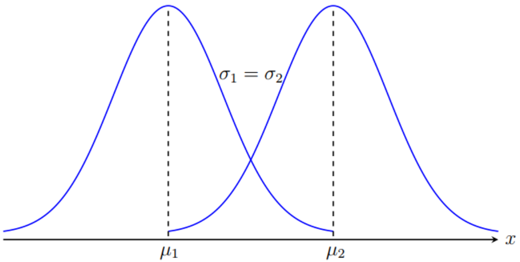
Định lý 11

Nếu X là biến ngẫu nhiên có phân phối chuẩn với các tham số μ và σ^2, thì

$$E(X) = \mu, \quad V(X) = \sigma^2 \quad \text{và} \quad \sigma_X = \sigma. \tag{45}$$

Kỳ vọng và phương sai

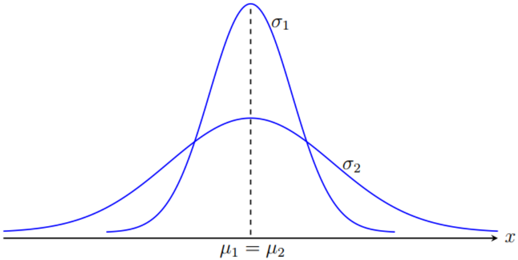
✍ Hình 19 mô tả hai đường cong chuẩn có cùng độ lệch chuẩn nhưng kỳ vọng khác nhau. Hai đường cong giống hệt nhau về hình dáng nhưng được tập trung tại các vị trí khác nhau dọc theo trục hoành.



Hình 19: Đường cong chuẩn với $\mu_1 < \mu_2$ và $\sigma_1 = \sigma_2$

Kỳ vọng và phương sai

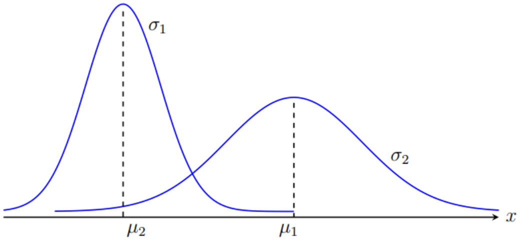
✍ Hình 20 mô tả hai đường cong chuẩn có cùng kỳ vọng nhưng độ lệch chuẩn khác nhau.



Hình 20: Đường cong chuẩn với $\mu_1 = \mu_2$ và $\sigma_1 < \sigma_2$

Kỳ vọng và phương sai

✍ Hình 21 mô tả hai đường cong chuẩn có kỳ vọng và độ lệch chuẩn khác nhau.



Hình 21: Đường cong chuẩn với $\mu_1 < \mu_2$ và $\sigma_1 < \sigma_2$

Kỳ vọng và phương sai

Định lý 12

Nếu X là biến ngẫu nhiên có phân phối chuẩn với tham số μ và σ^2 thì $Y = \alpha X + \beta$ là biến ngẫu nhiên có phân phối chuẩn $\mathcal{N}(\alpha\mu + \beta; \alpha^2\sigma^2)$, với α và β là các số thực tùy ý.

✍ Định lý 12 phát biểu rằng bất kỳ phép biến đổi tuyến tính nào của một biến ngẫu nhiên có phân phối chuẩn sẽ tạo ra một biến ngẫu nhiên có phân phối chuẩn khác.

Phân phối chuẩn tắc

✍ **Chuẩn hóa một biến ngẫu hiên phân phối chuẩn**

- Giả sử $X \sim \mathcal{N}(\mu; \sigma^2)$. Đặt

$$Z = \frac{X - \mu}{\sigma} \tag{46}$$

thì theo Định lý 12, Z là biến ngẫu nhiên có phân phối chuẩn với $E(Z) = 0$ và $V(Z) = 1$. Biến ngẫu nhiên Z này được gọi là biến ngẫu nhiên có phân phối chuẩn tắc.

- Phép biến đổi (46) được gọi là phép chuẩn tắc hóa một biến ngẫu nhiên $X \sim \mathcal{N}(\mu; \sigma^2)$.

Định nghĩa 18

Biến ngẫu nhiên Z có phân phối chuẩn $\mathcal{N}(0; 1)$ được gọi là biến ngẫu nhiên có phân phối chuẩn tắc.

Phân phối chuẩn tắc

Định nghĩa 19

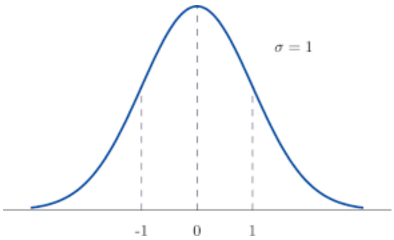
Hàm mật độ xác suất của biến ngẫu nhiên Z có phân phối chuẩn tắc $\mathcal{N}(0;1)$ là

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \ \forall z \in \mathbb{R}. \quad (47)$$

✍

- Hàm $\varphi(z)$ trong (47) là hàm Gauss với các giá trị được tính sẵn trong Bảng giá trị hàm Gauss.
- Đồ thị hàm $\varphi(z)$ cho trong Hình 22.

Phân phối chuẩn tắc



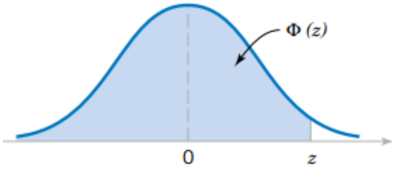
Hình 22: Hàm mật độ xác suất của biến ngẫu nhiên có phân phối chuẩn tắc $\mathcal{N}(0; 1)$

Phân phối chuẩn tắc

Định nghĩa 20

Hàm phân phối xác suất của biến ngẫu nhiên Z có phân phối chuẩn tắc là

$$\Phi(z) = P(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du \ \forall z \in \mathbb{R}. \quad (48)$$



Hình 23: $\Phi(z) = P(Z < z)$

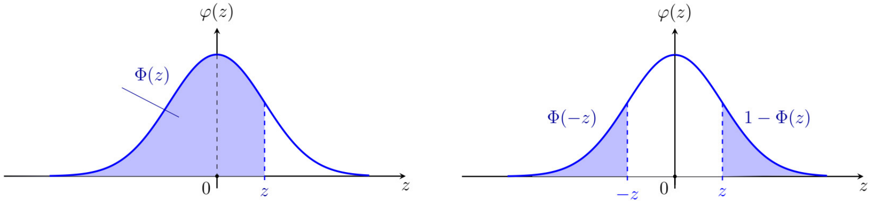
Phân phối chuẩn tắc

✍

- Giá trị của $\Phi(z)$ được tính sẵn trong bảng giá trị hàm $\Phi(z)$ với $z \geq 0$.
- Đối với các giá trị âm của z ta sử dụng tính chất (49) của hàm $\Phi(z)$.

Định lý 13

$$\Phi(-z) = 1 - \Phi(z). \tag{49}$$



Hình 24: Tính chất đối xứng của hàm mật độ xác suất của biến ngẫu nhiên phân phối chuẩn tắc

Phân phối chuẩn tắc

Ví dụ 41

Sử dụng bảng phân vị chuẩn, ta nhận được

- $\Phi(2,0) = 0,97725$;
- $\Phi(2,02) = 0,97831$;
- $\Phi(-2,0) = 1 - \Phi(2,0) = 1 - 0,97725 = 0,02275$.

(2)

(2) Trích bảng phân vị chuẩn $\Phi(z) = P(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du$

x	0	1	2	3	4	5	6	7	8	9
1,0	84134	84375	84614	84850	85083	85314	85543	85769	85993	86214
2,0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169

Phân phối chuẩn tắc

Định lý 14

Nếu X là biến ngẫu nhiên có phân phối chuẩn $\mathcal{N}(\mu;\sigma^2)$ và $Z = (X-\mu)/\sigma$ thì Z có phân phối chuẩn tắc $\mathcal{N}(0;1)$ và

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \tag{50}$$

trong đó, $F_X(.)$ là hàm phân phối xác suất của biến ngẫu nhiên X và $\Phi(.)$ là hàm phân phối xác suất của biến ngẫu nhiên Z.

$P(\alpha < X < \beta)$ với $X \sim \mathcal{N}(\mu;\sigma^2)$

✍

- Cho $X \sim \mathcal{N}(\mu;\sigma^2)$ và $\alpha, \beta \in \mathbb{R}$.
- Sử dụng tính chất của hàm phân phối xác suất và (50) ta nhận được

$$P(\alpha < X < \beta) = F_X(\beta) - F_X(\alpha) = \Phi\Big(\frac{\beta-\mu}{\sigma}\Big) - \Phi\Big(\frac{\alpha-\mu}{\sigma}\Big).$$

Định lý 15

Nếu X là biến ngẫu nhiên có phân phối chuẩn với các tham số μ và σ^2 thì

$$P(\alpha < X < \beta) = \Phi\Big(\frac{\beta-\mu}{\sigma}\Big) - \Phi\Big(\frac{\alpha-\mu}{\sigma}\Big), \quad \alpha, \beta \in \mathbb{R}, \tag{51}$$

trong đó, $\Phi(.)$ là hàm phân phối chuẩn tắc được định nghĩa trong (48).

$P(\alpha < X < \beta)$ với $X \sim \mathcal{N}(\mu; \sigma^2)$

SAMI 1956

Hệ quả 4

Nếu $X \sim \mathcal{N}(\mu; \sigma^2)$ thì

(a) $P(X < \beta) = \Phi\left(\dfrac{\beta - \mu}{\sigma}\right)$.

(b) $P(X > \alpha) = 1 - \Phi\left(\dfrac{\alpha - \mu}{\sigma}\right)$.

(c) $P(|X - \mu| < \varepsilon) = 2\Phi\big(\varepsilon/\sigma\big) - 1$.

SAMI 1956

Ví dụ 42

Giả sử phép đo cường độ dòng điện trên một sợi dây đồng là biến ngẫu nhiên có phân phối chuẩn với giá trị trung bình là 10 miliampe và phương sai là 4 $(\text{miliampe})^2$. Tính xác suất để nhận được kết quả của phép đo này vượt quá 13 miliampe.

Giải. Gọi X là phép đo cường độ dòng điện, $X \sim \mathcal{N}(\mu; \sigma^2)$ với $\mu = 10$ và $\sigma = 2$. Sử dụng Hệ quả 4(b),

$$P(X > 13) = 1 - \Phi\left(\frac{13-10}{2}\right) = 1 - \Phi(1,5) = 1 - 0,93319 = 0,06681.$$

(3)

(3) Trích bảng phân vị chuẩn $\Phi(z) = P(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du$

x	0	1	2	3	4	5	6	7	8	9
1,5	**93319**	93448	93574	93699	93822	93943	94062	94179	94295	94408
1,6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449

$P(\alpha < X < \beta)$ **với** $X \sim \mathcal{N}(\mu; \sigma^2)$

Ví dụ 43

Lãi suất (%) đầu tư vào một dự án trong năm 2022 được giả thiết là biến ngẫu nhiên có phân phối chuẩn. Theo đánh giá của ủy ban đầu tư thì với xác suất 0,15866 cho lãi suất lớn hơn 20% và với xác suất 0,02275 cho lãi suất lớn hơn 25%. Vậy khả năng đầu tư mà không bị lỗ là bao nhiêu?

Giải.

- Gọi X là lãi suất (%) của dự án trong năm 2022, $X \sim \mathcal{N}(\mu; \sigma^2)$.
- Theo đầu bài và Hệ quả 4(b),

$$P(X > 20) = 1 - \Phi\left(\frac{20-\mu}{\sigma}\right) = 0,15866 \quad \text{và} \quad P(X > 25) = 1 - \Phi\left(\frac{25-\mu}{\sigma}\right) = 0,02275.$$

$P(\alpha < X < \beta)$ **với** $X \sim \mathcal{N}(\mu;\sigma^2)$

- Hay

$$\Phi\Big(\frac{20-\mu}{\sigma}\Big) = 0,84134 \quad \text{và} \quad \Phi\Big(\frac{25-\mu}{\sigma}\Big) = 0,97725.$$

- Từ bảng phân vị chuẩn, $\frac{20-\mu}{\sigma} = 1$ và $\frac{25-\mu}{\sigma} = 2$. Suy ra $\mu = 15$, $\sigma = 5$.
- Vậy khả năng đầu tư không bị lỗ là

$$P(X \geq 0) = 1 - \Phi(\frac{0-15}{3}) = \Phi(3) = 0,99865.$$

(4)

---

(4) Trích bảng phân vị chuẩn $\Phi(z) = P(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du$

x	0	1	2	3	4	5	6	7	8	9
1,0	84134	84375	84614	84850	85083	85314	85543	85769	85993	86214
2,0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169
3,0	**0,99865**	3,1	99903	3,2	99931	3,3	99952	3,4	99966	

$P(\alpha < X < \beta)$ **với** $X \sim \mathcal{N}(\mu; \sigma^2)$

Nhận xét 2

Từ Hệ quả 4(c) suy ra một số kết quả hữu ích liên quan đến biến ngẫu nhiên X có phân phối chuẩn $\mathcal{N}(\mu; \sigma^2)$ được tóm tắt dưới đây.

$$P(\mu - \sigma < X < \mu + \sigma) = 0,68268,$$
$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0,9545,$$
$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0,9973.$$

Xấp xỉ phân phối nhị thức bởi phân phối chuẩn

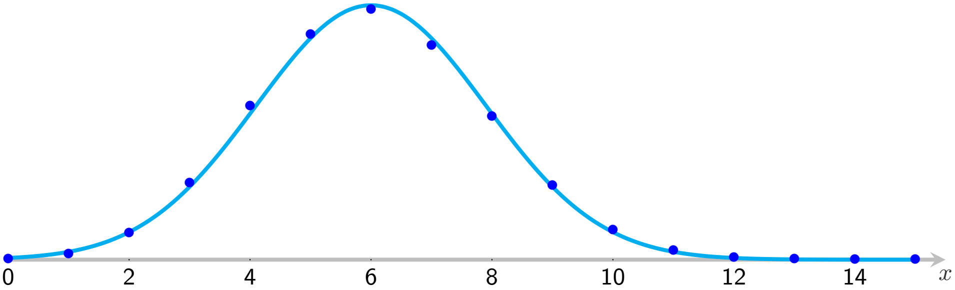
Định lý 16 (Định lý Moivre–Laplace)

Giả sử X là biến ngẫu nhiên có phân phối nhị thức với tham số n và p. Nếu $np \geq 5$ và $n(1-p) \geq 5$ thì X có phân phối xấp xỉ phân phối chuẩn với tham số $\mu = np$ và $\sigma^2 = np(1-p)$.

✍

- Phân phối chuẩn với kỳ vọng $\mu = np$ và phương sai $\sigma^2 = np(1-p)$ không chỉ xấp xỉ khá tốt cho phân phối nhị thức khi n khá lớn và xác suất p không quá gần 0 hoặc 1 mà còn cung cấp một xấp xỉ khá tốt cho phân phối nhị thức ngay cả khi n nhỏ và p gần $1/2$.
- Cho biến ngẫu nhiên $X \sim \mathcal{B}(n;p)$. Để minh họa việc xấp xỉ phân phối chuẩn cho phân phối nhị thức, ta vẽ phân phối nhị thức với $n = 15$, $p = 0,4$ và vẽ đường cong chuẩn có cùng kỳ vọng $\mu = np = 15 \times 0,4 = 6$ và cùng phương sai $\sigma^2 = np(1-p) = 15 \times 0,4 \times 0,6 = 3,6$ với X (Hình 25).

Xấp xỉ phân phối nhị thức bởi phân phối chuẩn



Hình 25: Xấp xỉ phân phối chuẩn cho phân phối nhị thức $\mathcal{B}(15; 0,4)$

Xấp xỉ phân phối nhị thức bởi phân phối chuẩn

SAMI 1956

Định lý 17 (Xấp xỉ phân phối nhị thức bằng phân phối chuẩn)

Cho X là biến ngẫu nhiên có phân phối nhị thức với tham số n và p. Phân phối xác suất của X được xấp xỉ bởi phân phối chuẩn với $\mu = np$ và $\sigma^2 = np(1-p)$ và

$$P(X = k) = P(k \leq X \leq k) \approx \Phi\left(\frac{k + 0,5 - \mu}{\sigma}\right) - \Phi\left(\frac{k - 0,5 - \mu}{\sigma}\right), \tag{52}$$

$$P(k_1 \leq X \leq k_2) \approx \Phi\left(\frac{k_2 + 0,5 - \mu}{\sigma}\right) - \Phi\left(\frac{k_1 - 0,5 - \mu}{\sigma}\right). \tag{53}$$

✍

- Vì ta xấp xỉ một phân phối rời rạc bằng một phân phối liên tục, nên cần một sự hiệu chỉnh để giảm sai số.
- Việc thêm $+0,5$ và $-0,5$ chính là yếu tố hiệu chỉnh và gọi là hiệu chỉnh liên tục.

Xấp xỉ phân phối nhị thức bởi phân phối chuẩn

Ví dụ 44

Sử dụng phân phối chuẩn xấp xỉ xác suất $P(8 \leq X \leq 10)$ cho biến ngẫu nhiên $X \sim \mathcal{B}(n;p)$ với $n = 25$ và $p = 0,5$. So sánh với công thức tính chính xác.

Giải.

- Vì $X \sim \mathcal{B}(n;p)$ với $n = 25$ và $p = 0,5$ nên

$$P(8 \leq X \leq 10) = \left(C_{25}^{8} + C_{25}^{9} + C_{25}^{10}\right) \times (0,5)^{25} \approx 0,190535.$$

- Sử dụng công thức xấp xỉ (53) với $\mu = np = 12,5$ và $\sigma = \sqrt{np(1-p)} = 2,5$ ta nhận được

$$\begin{aligned} P(8 \leq X \leq 10) &\approx \Phi\left(\frac{10 + 0,5 - 12,5}{2,5}\right) - \Phi\left(\frac{8 - 0,5 - 12,5}{2,5}\right) \\ &= \Phi(-0.8) - \Phi(-2) = \Phi(2) - \Phi(0,8) = 0,97725 - 0,78814 = 0,18911. \end{aligned}$$

- Giá trị xấp xỉ 0,18911 và giá trị chính xác 0,190535 là khá gần nhau.

e. Xấp xỉ phân phối nhị thức bởi phân phối chuẩn

Ví dụ 45

Kiểm tra chất lượng 1000 sản phẩm trong một lô hàng có tỷ lệ chính phẩm là 0,95. Tìm xác suất để trong 1000 sản phẩm đó có từ 940 đến 960 chính phẩm.

Giải.

- Gọi X là biến ngẫu nhiên chỉ số chính phẩm trong 1000 sản phẩm được kiểm tra, $X \sim \mathcal{B}(n;p)$ với $n = 1000$ và $p = 0,95$.
- Vì $np = 950$ và $n(1-p) = 50$ khá lớn, nên ta có thể xấp xỉ phân phối nhị thức $\mathcal{B}(n;p)$ bởi phân phối chuẩn $\mathcal{N}(\mu;\sigma^2)$ với $\mu = np = 950$, $\sigma^2 = np(1-p) = 47,5$ và

$$P(940 \leq X \leq 960) \approx \Phi\left(\frac{960+0,5-950}{\sqrt{47,5}}\right) - \Phi\left(\frac{940-0,5-950}{\sqrt{47,5}}\right)$$
$$= \Phi(1,52) - \Phi(-1,52) = 2\Phi(1,52) - 1 = 2 \times 0,93574 - 1 = 0,87148.$$

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối Khi-bình phương

Định nghĩa 21

Biến ngẫu nhiên liên tục X được gọi là có phân phối Khi-bình phương với ν bậc tự do, ký hiệu là $X \sim \chi^2(\nu)$, nếu hàm mật độ xác suất của nó có dạng:

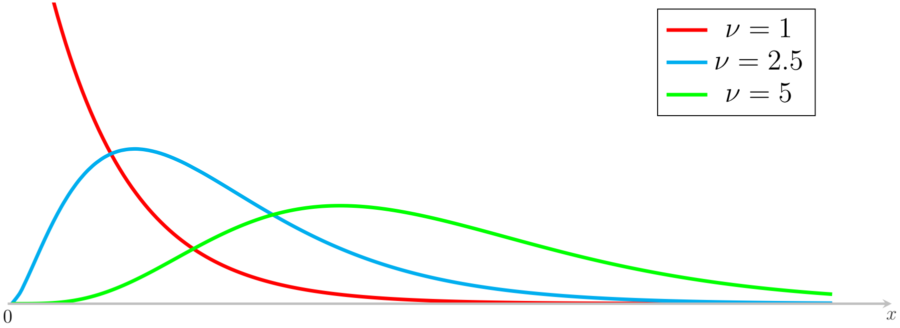
$$f_X(x) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2}, & x > 0, \\ 0, & x \le 0, \end{cases} \tag{54}$$

ở đây, $\Gamma(.)$ là hàm Gamma được định nghĩa bởi $\Gamma : (0; \infty) \to \mathbb{R}$,

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt. \tag{55}$$

✍ Đồ thị của phân phối Khi-bình phương bị lệch sang phải. Tuy nhiên, khi số bậc tự do ν tăng lên, phân phối này trở nên đối xứng hơn. Khi $\nu \to \infty$, dạng giới hạn của phân phối Khi-bình phương là phân phối chuẩn (Hình 26).

Định nghĩa phân phối Khi-bình phương



Hình 26: Hàm mật độ xác suất của phân phối Khi-bình phương với số bậc tự do khác nhau

Kỳ vọng, phương sai

Định lý 18

Nếu X là biến ngẫu nhiên có phân phối Khi-bình phương với ν bậc tự do, thì

$$E(X) = \nu \quad \text{và} \quad V(X) = 2\nu.$$

Giá trị tới hạn phân phối Khi-bình phương

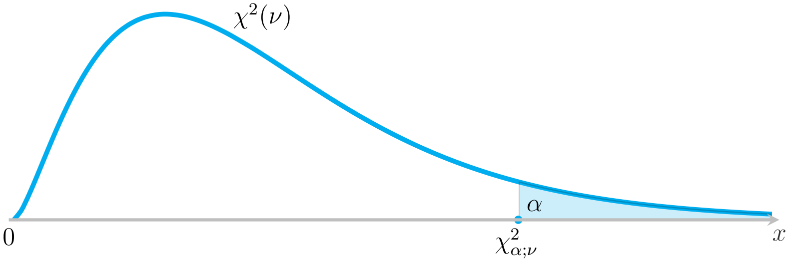
Giá trị tới hạn $\chi^2_{\alpha;\nu}$

- Ký hiệu $\chi^2_{\alpha;\nu}$ chỉ giá trị của biến ngẫu nhiên $X \sim \chi^2(\nu)$ sao cho xác suất X vượt quá giá trị này là α. Nghĩa là,

$$P(X > \chi^2_{\alpha;\nu}) = \int_{\chi^2_{\alpha,\nu}}^{\infty} f_X(x)dx = \alpha \quad \text{với } f_X(x) \text{ xác định bởi (54).}$$

- Xác suất này là miền bóng mờ trong Hình 27.

Giá trị tới hạn phân phối Khi-bình phương



Hình 27: $P(X > \chi^2_{\alpha;\nu}) = \alpha$

Giá trị tới hạn phân phối Khi-bình phương

✍

- Giá trị $\chi^2_{\alpha;\nu}$ được tính sẵn trong Bảng giá trị tới hạn phân phối Khi-bình phương.
- Chẳng hạn, giá trị với 9 bậc tự do có xác suất ở bên phải bằng $0,025$ là $\chi^2_{0,025;9} = 19,023$, tức là

$$P(X > \chi^2_{0,025;9}) = P(X > 19,023) = 0,025.$$

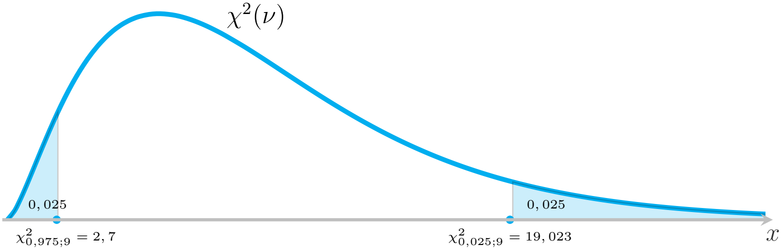
- Ngược lại, ta có $\chi^2_{0,975;9} = 2,700$ (Hình 28).

(5)

(5) Trích Bảng giá trị tới hạn phân phối Khi-bình phương

Bậc tự do \ α	$0,995$	$0,990$	$0,975$	$0,95$	$0,050$	$0,025$	$0,010$	0,005
9	1,735	2,088	2,700	3,325	16,919	19,023	21,666	23,590
10	2,156	2,558	3,247	3,940	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	19,675	21,920	24,725	26,758

Giá trị tới hạn phân phối Khi-bình phương



Hình 28: $\chi^2_{0,025;9} = 19,023$ và $\chi^2_{0,975;9} = 2,700$

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối Student

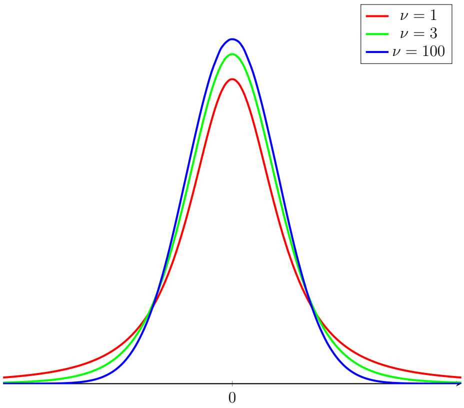
Định nghĩa 22

Biến ngẫu nhiên liên tục T được gọi là có phân phối Student (hay phân phối t) với ν bậc tự do, ký hiệu là $T \sim t(\nu)$, nếu hàm mật độ xác suất của T có dạng

$$f_T(t) = \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}, \quad -\infty < t < +\infty, \tag{56}$$

ở đây, $\Gamma(x)$ là hàm Gamma định nghĩa bởi (55).

Định nghĩa phân phối Student



Hình 29: Hàm mật độ xác suất của phân phối Student với số bậc tự do khác nhau

Định nghĩa phân phối Student

✍

- Phân phối Student có cùng dạng và tính đối xứng như phân phối chuẩn nhưng nó phản ánh tính biến đổi của phân phối sâu sắc hơn.
- Khi số bậc tự do ν tăng lên thì phân phối Student tiến nhanh về phân phối chuẩn. Trong thực hành, khi $\nu \geq 30$ ta có thể dùng phân phối chuẩn thay thế cho phân phối Student.

Định lý 19

Cho Z là biến ngẫu nhiên có phân phối chuẩn tắc và V là biến ngẫu nhiên có phân phối Khi-bình phương với ν bậc tự do. Nếu Z và V là hai biến ngẫu nhiên độc lập thì thống kê

$$T = \frac{Z}{\sqrt{\frac{V}{n}}} \tag{57}$$

có phân phối Student với ν bậc tự do.

Kỳ vọng, phương sai

Định lý 20

Nếu X là biến ngẫu nhiên phân phối Student với ν bậc tự do, thì

$$E(X) = 0 \quad \text{và} \quad V(X) = \frac{\nu}{\nu - 2} \quad \text{với} \quad \nu > 2.$$

Giá trị tới hạn phân phối Student

Giá trị tới hạn $t_{\alpha;\nu}$

- Ký hiệu $t_{\alpha;\nu}$ chỉ giá trị của biến ngẫu nhiên $T \sim t(\nu)$ sao cho xác suất T vượt quá giá trị này là α. Nghĩa là,

$$P(T > t_{\alpha;\nu}) = \alpha.$$

- Giá trị $t_{\alpha;\nu}$ được tính sẵn trong Bảng giá trị tới hạn phân phối Student.
- Chẳng hạn, giá trị với 10 bậc tự do có xác suất ở bên phải bằng $0,05$ là $t_{0,05;10} = 1,812$, tức là

$$P(T > t_{0,05;10}) = P(T > 1,812) = 0,05.$$

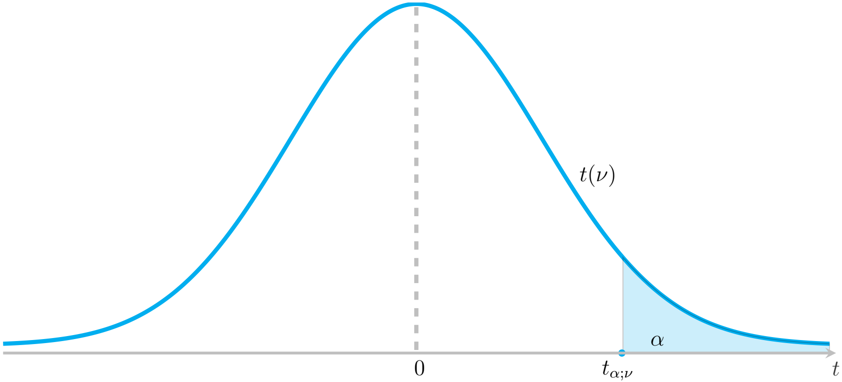
- Vì tính đối xứng nên $t_{0,95;10} = -t_{0,05;10} = -1,812$.

(6)

[6] Trích Bảng giá trị tới hạn phân phối Student

Bậc tự do \ α	$0,10$	$0,05$	$0,025$	$0,01$	$0,005$	$0,0025$	$0,001$	$0,0005$
9	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	1,372	**1,812**	2,228	2,764	3,169	3,581	4,144	4,587

Giá trị tới hạn phân phối Student



Hình 30: $P(T > t_{\alpha;\nu}) = \alpha$

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Định nghĩa phân phối Fisher

Định nghĩa 23

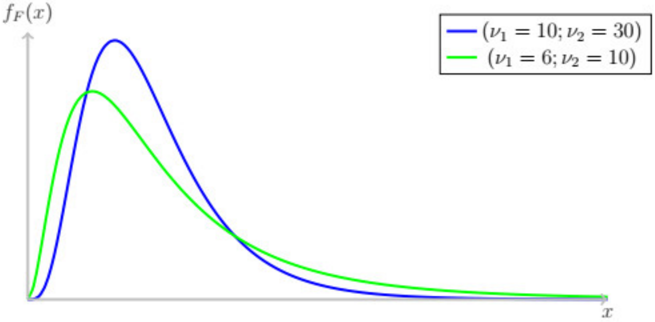
Biến ngẫu nhiên liên tục X được gọi là có phân phối Fisher (hay phân phối F) với ν_1 và ν_2 bậc tự do, ký hiệu là $X \sim F(\nu_1;\nu_2)$, nếu hàm mật độ xác suất của nó có dạng

$$f_X(x) = \begin{cases} \frac{\Gamma[(\nu_1+\nu_2)/2](\nu_1/\nu_2)^{\nu_1/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \frac{x^{\nu_1/2-1}}{(1+\nu_1 x/\nu_2)^{(\nu_1+\nu_2)/2}}, & x > 0, \\ 0, & x \leq 0, \end{cases} \tag{58}$$

ở đây $\Gamma(.)$ là hàm Gamma được định nghĩa bởi (55).

✍ Hàm mật độ xác suất (58) không chỉ phụ thuộc vào hai tham số ν_1 và ν_2 mà còn phụ thuộc vào thứ tự xuất hiện của chúng. Hình 31 biểu thị một số phân phối Fisher cơ bản.

Định nghĩa phân phối Fisher



Hình 31: Một số phân phối Fisher cơ bản

Định nghĩa phân phối Fisher

Định lý 21

Cho U và V là hai biến ngẫu nhiên độc lập có phân phối Khi-bình phương với ν_1 và ν_2 bậc tự do tương ứng. Khi đó, biến ngẫu nhiên

$$F = \frac{U/\nu_1}{V/\nu_2}, \tag{59}$$

có phân phối Fisher với ν_1 và ν_2 bậc tự do.

Kỳ vọng, phương sai

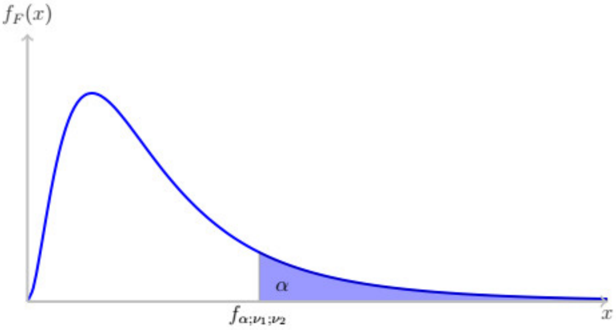
Định lý 22

Nếu X là biến ngẫu nhiên có phân phối Fisher với ν_1 và ν_2 bậc tự do, thì

$$E(X) = \frac{\nu_1}{\nu_1 - 2},\ \nu_1 > 2 \quad \text{và} \quad V(X) = \frac{2\nu_2^2(\nu_1 + \nu_2^2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)},\ \nu_2 > 4. \tag{60}$$

Giá trị tới hạn phân phối Fisher

Ký hiệu $F_{\alpha;\nu_1;\nu_2}$ là giá trị tới hạn phân phối Fisher ứng với bậc tự do ν_1 và ν_2 mà ở đó ta tìm được diện tích của miền phẳng được giới hạn bởi đường cong $f_X(x)$, trục hoành và bên phải của $F_{\alpha;\nu_1;\nu_2}$ bằng α. Ta minh họa điều này bằng hình bóng mờ trong Hình 32.



Hình 32: Giá trị tới hạn $f_{\alpha;\nu_1;\nu_2}$

Giá trị tới hạn phân phối Fisher

✍

- Giá trị $F_{\alpha;\nu_1;\nu_2}$ với $\alpha = 0,05$ và $\alpha = 0,01$ và với các tổ hợp khác nhau của bậc tự do ν_1 và ν_2 được tính sẵn trong Bảng giá trị tới hạn phân phối Fisher.
- Chẳng hạn, giá trị ở bậc tự do $\nu_1 = 6$ và $\nu_2 = 10$ cho diện tích 0,05 ở bên phải là $F_{0,05;6;10} = 3,22$.
- Định lý dưới đây giúp ta tìm được $F_{0,95;\nu_1;\nu_2}$ và $F_{0,99;\nu_1;\nu_2}$.

(7)

---

(7) Trích Bảng giá trị tới hạn phân phối Fisher với $\alpha = 0,05$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
10	4,96	4,10	3,71	3,48	3,33	**3,22**	3,14	3,07	3,02

Giá trị tới hạn phân phối Fisher

Định lý 23

Ta có

$$F_{1-\alpha;\nu_1;\nu_2} = \frac{1}{F_{\alpha;\nu_2;\nu_1}}. \tag{61}$$

Ví dụ 46

$$F_{0,95;6;10} = \frac{1}{F_{0,05;10;6}} = \frac{1}{4,06} = 0,2463.$$

(8)

---

(8) Trích Bảng giá trị tới hạn phân phối Fisher với $\alpha = 0,05$

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
5	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	**4,06**	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67

2.4. MỘT SỐ PHÂN PHỐI XÁC SUẤT THÔNG DỤNG

Bài tập

SAMI 1956

Bài 18

Cho X là biến ngẫu nhiên có phân phối nhị thức $X \sim \mathcal{B}(n;p)$. Khẳng định nào sau đây là ĐÚNG?

A. $E(X) = np(1-p)$.

B. $V(X) = np$.

C. $\text{mod}(X) = [np + 1 - p]$.

D. $\sigma(X) = \sqrt{np(1-p)}$.

Bài 19

Tỷ lệ sản phẩm lỗi trong một dây chuyền sản xuất là 0,03. Chọn ra 500 sản phẩm do dây chuyền này sản xuất để kiểm tra và ký hiệu X là số sản phẩm đạt yêu cầu trong 500 sản phẩm đó. Khẳng định nào sau đây SAI?

A. $X \sim \mathcal{B}(500; 0,03)$.

B. $X \sim \mathcal{B}(500; 0,97)$.

C. $E(X) = 485$.

D. $V(X) = 14,55$.

Bài tập

Bài 20

Tỷ lệ người có nhóm máu O tại một địa phương là 0,3. Hỏi phải chọn ít nhất bao nhiêu người để với xác suất không nhỏ hơn 0,95 có thể tin rằng có ít nhất một người có nhóm máu O?

Bài 21

Giả sử số khách hàng đến mua hàng tại một siêu thị là biến ngẫu nhiên có phân phối Poisson. Biết rằng trung bình mỗi giờ có 300 khách hàng đến mua hàng tại siêu thị này.

(a) Tính xác suất để có đúng 2 khách hàng đến mua hàng tại siêu thị đó trong một phút.

(b) Tính xác suất để có đúng 5 khách hàng đến mua hàng tại siêu thị đó trong ba phút.

(c) Tính xác suất để trong 3 phút liên tiếp mỗi phút có ít nhất một khách hàng đến mua hàng tại siêu thị.

Bài tập

Bài 22

Giả sử số cuộc gọi bảo hành một loại sản phẩm gia dụng đến trung tâm bảo hành là biến ngẫu nhiên có phân phối Poisson. Kinh nghiệm thấy rằng trung bình cứ mỗi giờ lại có 20 cuộc gọi đến trung tâm bảo hành. Trong một ngày làm việc, nếu như trong 5 phút đầu tiên khi trung tâm bắt đầu nhận cuộc gọi không có cuộc gọi nào đến thì xác suất để có cuộc gọi trong 5 phút tiếp theo là bao nhiêu?

Bài 23

Một thiết bị điện tử bao gồm 3000 bóng bán dẫn hoạt động độc lập với nhau. Nhà sản xuất tuyên bố rằng khả năng để mỗi bóng bị hỏng là 0,0003. Thiết bị điện tử sẽ chỉ hoạt động khi không có bóng nào bị hỏng. Tính xác suất để thiết bị điện tử đó ngừng hoạt động.

Bài tập

Bài 24

Một thiết bị điện tử bao gồm 10000 bóng bán dẫn hoạt động độc lấp với nhau, trong đó có 1000 bóng loại I, 3000 bóng loại II và 6000 bóng loại III. Xác xuất để bóng loại I, loại II và loại III bị hỏng lần lượt là 0,0005; 0,0003 và 0,0001. Thiết bị điện tử sẽ ngừng hoạt động nếu có ít nhất một bóng bị hỏng. Tính xác suất để thiết bị điện tử ngừng hoạt động.

Bài 25

Cho biến ngẫu nhiên X có phân phối đều trên $[0; 1]$. Tính xác suất để trong 10 lần quan sát thấy có 8 lần X nhận giá trị trong khoảng $(0, 2; 0, 8)$.

Bài tập

Bài 26

Tuyến xe bus Hà Nội – Hưng Yên khởi hành từ bến xe Giáp Bát lúc 6h sáng và cứ 30 phút lại có một chuyến xuất bến. Một khách hàng đến bến xe trong khoảng thời gian từ 7 giờ đến 8 giờ. Tính xác suất để người đó phải chờ trên 15 phút.

Bài 27

Ký hiệu X (kg) là trọng lượng của một bao bột mì được đóng gói từ một dây chuyền tự động. Giả sử X là biến ngẫu nhiên có phân phối chuẩn, $X \sim \mathcal{N}(15; 0,2^2)$.

(a) Tính tỷ lệ các bao bột mì có trọng lượng sai lệch so với trọng lượng tiêu chuẩn 15 kg không quá 400 g.

(b) Chọn ngẫu nhiên một bao bột từ dây chuyền sản xuất đó. Tính xác suất để bao được chọn có trọng lượng trên 15,4 kg.

(c) Khi chọn ra 100 bao thì số bao có khả năng nhất có trọng lượng sai lệch so với trọng lượng tiêu chuẩn 15 kg không quá 400 g là bao nhiêu?

Bài tập

Bài 28

Tính xác suất để trong 1000 sản phẩm có không quá 16 sản phẩm bị lỗi, biết tỷ lệ sản phẩm lỗi trong dây chuyển sản xuất đó là 0,03.

Bài 29

Trọng lượng tịnh của một loại sữa hộp là 1 kg với sai số cho phép là 0,05 kg. Nếu hộp sữa có trọng lượng dưới 0,95 kg thì được coi là không đạt tiêu chuẩn. Trong một dây chuyền sản xuất, người ta ước lượng rằng tỷ lệ hộp sữa không đạt yêu cầu về trọng lượng là 0,03. Kiểm tra ngẫu nhiên trọng lượng của 1000 hộp sữa. Tính xác suất để có trên 950 hộp đạt tiêu chuẩn về trọng lượng.

Bài tập

Bài 30

Biết rằng tuổi thọ X (giờ) của một loại đèn compact là biến ngẫu nhiên có phân phối mũ với tham số $\lambda = 0,0005$.

(a) Tuổi thọ trung bình của loại bóng đèn trên là bao nhiêu?

(b) Tính xác suất để bóng đèn có thể chiếu sáng được ít nhất 1000 giờ.

(c) Nếu biết bóng đèn đã chiếu sáng được 500 giờ thì xác suất để bóng đèn chiếu sáng thêm 500 giờ nữa là bao nhiêu?