Chương 2

ĐẠI LƯỢNG NGẪU NHIÊN VÀ PHÂN PHỐI XÁC SUẤT

1. ĐẠI LƯỢNG NGẪU NHIÊN

1.1 Khái niệm đại lượng ngẫu nhiên

□ **Định nghĩa 1** *Đại lượng ngẫu nhiên là đại lượng biến đổi biểu thị giá trị kết qủa của một phép thử ngẫu nhiên.*

Ta dùng các chữ cái hoa như X, Y, Z, ... để kí hiệu đại lượng ngẫu nhiên.

• **Ví dụ 1** *Tung một con xúc xắc. Gọi X là số chấm xuất hiện trên mặt con xúc xắc thì X là một đại lượng ngẫu nhiên nhận các giá trị có thể là 1, 2, 3, 4, 5, 6.*

1.2 Đại lượng ngẫu nhiên rời rạc

a) Đại lượng ngẫu nhiên rời rạc

□ **Định nghĩa 2** *Đại lượng ngẫu nhiên được gọi là rời rạc nếu nó chỉ nhận một số hữu hạn hoặc một số vô hạn đếm được các giá trị.*

Ta có thể liệt kê các giá trị của đại lượng ngẫu nhiên rời rạc $x_1, x_2, \ldots, x_n$.

Ta kí hiệu đại lượng ngẫu nhiên X nhận giá trị x_n là $X = x_n$ và xác suất để X nhận giá trị x_n là $P(X = x_n)$.

• **Ví dụ 2** *Số chấm xuất hiện trên mặt con xúc xắc, số học sinh vắng mặt trong một buổi học...là các đại lượng ngẫu nhiên rời rạc.*

b) Bảng phân phối xác suất

Bảng phân phối xác suất dùng để thiết lập luật phân phối xác suất của đại lượng ngẫu nhiên rời rạc, nó gồm 2 hàng: hàng thứ nhất liệt kê các giá trị có thể $x_1, x_2, \ldots, x_n$ của đại lượng ngẫu nhiên X và hàng thứ hai liệt kê các xác suất tương ứng $p_1, p_2, \ldots, p_n$ của các giá trị có thể đó.

X	x_1	x_2	$\ldots$	x_n
P	p_1	p_2	$\ldots$	p_n

Nếu các giá trị có thể của đại lượng ngẫu nhiên X gồm hữu hạn số $x_1, x_2, \ldots, x_n$ thì các biến cố $X = x_1, X = x_2, \ldots, X = x_n$ lập thành một nhóm các biến cố đầy đủ xung khắc từng đôi.

Do đó $\sum_{i=1}^{n} p_i = 1$.

• **Ví dụ 3** *Tung một con xúc xắc đồng chất. Gọi X là số chấm xuất hiện trên mặt con xúc xắc thì X là đại lượng ngẫu nhiên rời rạc có phân phối xác suất cho bởi:*

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

1.3 Đại lượng ngẫu nhiên liên tục và hàm mật độ xác suất

a) Đại lượng ngẫu nhiên liên tục

□ **Định nghĩa 3** *Đại lượng ngẫu nhiên được gọi là liên tục nếu các giá trị có thể của nó lấp đầy một khoảng trên trục số.*

• **Ví dụ 4**

- *Nhiệt độ không khí ở mỗi thời điểm nào đó.*
- *Sai số khi khi đo lường một đại lượng vật lý.*
- *Khoảng thời gian giữa hai ca cấp cứu của một bệnh viện.*

b) Hàm mật độ xác suất

□ **Định nghĩa 4** *Hàm mật độ xác suất của đại lượng ngẫu nhiên liên tục X là hàm không âm f(x), xác định với mọi $x \in (-\infty, +\infty)$ thỏa mãn*

$$P(X \in B) = \int_B f(x)dx$$

với mọi tập số thực B.

◇ **Tính chất** Hàm mật độ xác suất có các tính chất sau

i) $f(x) \geq 0, \ \forall x \in (-\infty, +\infty)$

ii) $\int_{-\infty}^{+\infty} f(x)dx = 1$

⊙ **Ý nghĩa của hàm mật độ**

Từ định nghĩa của hàm mật độ ta có $P(x \leq X \leq x + \triangle x) \sim f(x).\triangle x$

Do đó ta thấy xác suất để X nhận giá trị thuộc lân cận khá bé $(x, x + \triangle x)$ gần như tỉ lệ với f(x).

1.4 Hàm phân phối xác suất

□ **Định nghĩa 5** *Hàm phân phối xác suất của đại lượng ngẫu nhiên X, kí hiệu F(x), là hàm được xác định như sau*

$$F(x) = P(X < x)$$

** Nếu X là đại lượng ngẫu nhiên rời rạc nhận các giá trị có thể* $x_1, x_2, \ldots, x_n$ *thì*

$$F(x) = \sum_{x_i<x} P(X = x_i) = \sum_{x_i<x} p_i \qquad (\textit{với } p_i = P(X = x_i))$$

** Nếu X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất f(x) thì*

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

◇ **Tính chất** Ta có thể chứng minh được các công thức sau

i) $0 \le F(x) \le 1; \quad \forall x.$

ii) F(x) là hàm không giảm ($x_1 \le x_2 \Longrightarrow F(x_1) \le F(x_2)$).

iii) $\lim\limits_{x\to-\infty} F(x) = 0; \quad \lim\limits_{x\to+\infty} F(x) = 1.$

iv) $F'(x) = f(x)$, $\forall x$.

⊙ **Ý nghĩa của hàm phân phối xác suất**

Hàm phân phối xác suất F(x) phản ánh mức độ tập trung xác suất về bên trái của điểm x.

• **Ví dụ 5** *Cho đại lượng ngẫu nhiên rời rạc X có bảng phân phối xác suất*

X	1	3	6
P	0,3	0,1	0,6

Tìm hàm phân phối xác suất của X và vẽ đồ thị của hàm này.

Giải

Nếu $x \le 1$ thì $F(x) = 0$.

Nếu $1 < x \le 3$ thì $F(x) = 0,3$.

Nếu $3 < x \le 6$ thì $F(x) = 0,3 + 0,1 = 0,4$.

Nếu $x > 6$ thì $F(x) = 0,3 + 0,1 + 0,6 = 1$.

$$F(x) = \begin{cases} 0 & ; \quad x \le 1 \\ 0,3 & ; \quad 1 < x \le 3 \\ 0,4 & ; \quad 3 < x \le 6 \\ 1 & ; \quad x > 6 \end{cases}$$

• **Ví dụ 6** *Cho X là đại lượng ngẫu nhiên liên tục có hàm mật độ*

$$f(x) = \begin{cases} 0 & \text{nếu} \quad x < 0 \\ \frac{6}{5}x & \text{nếu} \quad 0 \le x \le 1 \\ \frac{6}{5x^4} & \text{nếu} \quad x > 1 \end{cases}$$

Tìm hàm phân phối xác suất F(x).

Giải

Khi $x < 0$ thì $F(x) = \int\limits_{-\infty}^{x} f(t)dt = 0$

Khi $0 \le x \le 1$ thì $F(x) = \int\limits_{-\infty}^{x} f(t)dt = \int\limits_{0}^{x} \frac{6}{5}tdt = \frac{3}{5}x^2.$

Khi $x > 1$ thì

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt = \int\limits_{0}^{1} \frac{6}{5}tdt + \int\limits_{1}^{x} \frac{6}{5t^4}dt = \frac{3}{5} + \left[-\frac{2}{5t^3}\right]_1^x = 1 - \frac{2}{5x^3}$$

Vậy $F(x) = \begin{cases} 0 & ; \quad x < 0 \\ \frac{3}{5}x^2 & ; \quad 0 \le x \le 1 \\ 1 - \frac{2}{5x^3} & ; \quad x > 1 \end{cases}$

2. CÁC THAM SỐ ĐẶC TRƯNG CỦA ĐẠI LƯỢNG NGẪU NHIÊN

2.1 Kỳ vọng (Expectation)

□ **Định nghĩa 6**

** Giả sử X là đại lượng ngẫu nhiên rời rạc có thể nhận các giá trị $x_1, x_2, \ldots, x_n$ với các xáx suất tương ứng $p_1, p_2, \ldots, p_n$. Kỳ vọng của đại lượng ngẫu nhiên X, kí hiệu E(X) (hay M(X)), là số được xác định bởi*

$$E(X) = \sum_{i=1}^{n} x_i p_i$$

** Giả sử X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$. Kỳ vọng của đại lượng ngẫu nhiên X được xác định bởi*

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

• **Ví dụ 7** *Tìm kỳ vọng của đại lượng ngẫu nhiên có bảng phân phối xác suất sau*

X	5	6	7	8	9	10	11
P	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Ta có

$E(X) = 5.\frac{1}{12} + 6.\frac{2}{12} + 7.\frac{3}{12} + 8.\frac{2}{12} + 9.\frac{2}{12} + 10.\frac{1}{12} + 11.\frac{1}{12} = \frac{93}{12} = \frac{31}{4} = 7,75.$

• **Ví dụ 8** *Cho X là đại lượng ngẫu nhiên liên tục có hàm mật độ*

$$f(x) = \begin{cases} 2.e^{-2x} & \text{nếu } \; 0 < x < 2 \\ 0 & \text{nếu } \; x \notin (0,2) \end{cases}$$

Tìm E(X).

Giải

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{0}^{2} x.(\frac{1}{2}x) dx = \left.\frac{x^3}{6}\right|_0^2 = \frac{4}{3}$$

◇ **Tính chất**

i) $E(C) = C$, C là hằng.

ii) $E(cX) = c.E(X)$.

iii) $E(X + Y) = E(X) + E(Y)$.

iv) Nếu X và Y là hai đại lượng ngẫu nhiên độc lập thì $E(XY) = E(X).E(Y)$.

⊙ **Ý nghĩa của kỳ vọng**

Tiến hành n phép thử. Giả sử X là đại lượng ngẫu nhiên nhận các giá trị có thể $x_1, x_2, \ldots, x_n$ với số lần nhận $k_1, k_2, \ldots, k_n$.

Giá trị trung bình của đại lượng ngẫu nhiên X trong n phép thử là

$$\overline{x} = \frac{k_1 x_1 + k_2 x_2 + \ldots + k_n x_n}{n} = \frac{k_1}{x} x_1 + \frac{k_2}{n} x_2 + \ldots + \frac{k_n}{n} x_n = f_1 x_1 + f_2 x_2 + \ldots + f_n k_n$$

với $f_i = \frac{k_i}{n}$ là tần suất để X nhận giá trị x_i.

Theo định nghĩa xác suất theo lối thống kê ta có $\lim\limits_{n\to\infty} f_i = p_i$. Vì vậy với n đủ lớn ta có

$$\overline{x} \approx p_1x_1 + p_2x_2 + \ldots + p_nx_n = E(X)$$

Ta thấy kỳ vọng của đại lượng ngẫu nhiên xấp xỉ với trung bình số học các giá trị quan sát của đại lượng ngẫu nhiên.

Do đó có thể nói *kỳ vọng của đại lượng ngẫu nhiên chính là giá trị trung bình (theo xác suất) của đại lượng ngẫu nhiên. Nó phản ánh giá trị trung tâm của phân phối xác suất*

2.2 Phương sai (Variance)

□ **Định nghĩa 7** *Phương sai (độ lệch bình phương trung bình) của đại lượng ngẫu nhiên X, kí hiệu Var(X) hay D(X), được định nghĩa bằng công thức*

$$Var(X) = E\{[X - E(X)]^2\}$$

** Nếu X là đại lượng ngẫu nhiên rời rạc nhận các giá trị có thể $x_1, x_2, \ldots, x_n$ với các xác suất tương ứng $p_1, p_2, \ldots, p_n$ thì*

$$Var(X) = \sum_{i=1}^{n} [x_i - E(X)]^2 p_i$$

** Nếu X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất f(x) thì*

$$Var(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)dx$$

⊙ **Chú ý** Trong thực tế ta thường tính phương sai bằng công thức

$$\boxed{Var(X) = E(X^2) - [E(X)]^2}$$

Thật vậy, ta có

$$\begin{aligned} Var(X) &= E\{X - E(X)]^2\} \\ &= E\{X^2 - 2X.E(X) + [E(X)]^2\} \\ &= E(X^2) - 2E(X).E(X) + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

• **Ví dụ 9** *Cho đại lượng ngẫu nhiên rời rạc X có bảng phân phối xác suất sau*

X	1	3	5
P	0,1	0,4	0,5

Tìm phương sai của X.

Giải

E(X)=1.0,1+3.0,4+5.0,5=3,8

$E(X^2) = 1^2.0,1 + 3^2.0,4 + 5^2.0,5 = 16,2$

Do đó $Var(X) = E(X^2) - [E(X)]^2 = 16,2 - 14,44 = 1,76$.

• **Ví dụ 10** *Cho đại lượng ngẫunhiên X có hàm mật độ*

$$f(x) = \begin{cases} cx^3 & \text{với } 0 \le x \le 3 \\ 0 & \text{với } x \notin [0,3] \end{cases}$$

Hãy tìm

i) Hằng số c.

ii) Kỳ vọng.

iii) Phương sai

Giải

i) Ta có $1 = \int\limits_0^3 cx^3dx = c\left[\frac{x^4}{4}\right]_0^3 = \frac{81}{4}c.$

Suy ra $c = \frac{4}{81}$.

ii) $E(X) = \int\limits_0^3 x\frac{4}{81}x^3dx = \frac{4}{81}\left[\frac{x^5}{5}\right]_0^3 = 2,4.$

iii) Ta có

$$E(X^2) = \int\limits_{-\infty}^{\infty} x^2 f(x)dx = \int\limits_0^3 x^2\frac{4}{81}x^3dx = \frac{4}{81}\left[\frac{x^6}{6}\right]_0^3 = 6$$

Vậy $Var(X) = E(X^2) - [E(X)]^2 = 6 - (2,4)^2 = 0,24.$

◇ **Tính chất**

i) Var(C)=0; (C không đổi).

ii) $Var(cX) = c^2.Var(X)$.

iii) Nếu X và Y là hai đại lượng ngẫu nhiên độc lập thì

* $Var(X+Y) = Var(X) + Var(Y)$;

* Var(X-Y)=Var(X)+Var(Y);

* Var(C+X)=Var(X).

⊙ **Ý nghĩa của phương sai**

Ta thấy $X - E(X)$ là độ lệch khỏi giá trị trung bình nên $Var(X) = E\{[X - E(X)]^2\}$ là độ lệch bình phương trung bình. Do đó phương sai phản ánh mức độ phân tán các giá trị của đại lượng ngẫu nhiên chung quanh giá trị trung bình.

2.3 Độ lệch tiêu chuẩn

Đơn vị đo của phương sai bằng bình phương đơn vị đo của đại lượng ngẫu nhiên. Khi cần đánh giá mức độ phân tán các giá trị của đại lương ngẫu nhiên theo đơn vị của nó, người ta dùng một đặc trưng mới đó là độ lệch tiêu chuẩn.

□ **Định nghĩa 8** *Độ lệch tiêu chuẩn của đại lượng ngẫu nhiên* X*, kí hiệu là* $\sigma(X)$*, được định nghĩa như sau:*

$$\sigma(X) = \sqrt{Var(X)}$$

2.4 Mode

□ **Định nghĩa 9** *Mod(X) là giá trị của đại lượng ngẫu nhiên X có khả năng xuất hiện lớn nhất trong một lân cận nào đó của nó.*

Đối với đại lượng ngẫu nhiên rời rạc mod(X) là giá trị của X ứng với xác suất lớn nhất, còn đối với đại lượng ngẫu nhiên liên tục thì mod(X) là giá trị của X tại đó hàm mật độ đạt giá trị cực đại.

⊙ **Chú ý** Một đại lượng ngẫu nhiên có thể có một mode hoặc nhiều mode.

• **Ví dụ 11** *Giả sử X là điểm trung bình của sinh viên trong trường thì mod(X) là điểm mà nhiều sinh viên đạt được nhất.*

• **Ví dụ 12** *Cho đại lượng ngẫu nhiên liên tục có phân phối Vây−bun với hàm mật độ*

$$f(x) = \begin{cases} 0 & \textit{nếu} \quad x \leq 0 \\ \frac{x}{2}e^{-\frac{x^2}{4}} & \textit{nếu} \quad x > 0 \end{cases}$$

Hãy xác định mod(X).

Giải

mod(X) là nghiệm của phương trình

$$f'(x) = \frac{1}{2}e^{-\frac{x^2}{4}} - \frac{x^2}{4}e^{-\frac{x^2}{4}} = 0$$

Suy ra mod(X) là nghiệm của phương trình $1 - \dfrac{x^2}{2} = 0$. Do $mod(X) > 0$ nên $mod(X) = \sqrt{2} = 1,414$.

2.5 Trung vị

□ **Định nghĩa 10** *Trung vị của đại lượng ngẫu nhiên X là giá trị của X chia phân phối xác suất thành hai phần có xác suất giống nhau. Kí hiệu med(X).*

Ta có $P(X < med(X)) = P(X \geq med(X)) = \frac{1}{2}$

⊕ **Nhận xét** Từ định nghĩa ta thấy để tìm trung vị chỉ cần giải phương trình $F(x) = \frac{1}{2}$. Trong ứng dụng, trung vị là đặc trưng vị trí tốt nhất, nhiều khi tốt hơn cả kỳ vọng, nhất là khi trong số liệu có nhiều sai sót. Trung vị còn được gọi là *phân vị 50% của phân phối.*

• **Ví dụ 13** *Tìm med(X) trong ví dụ (12).*

Giải

med(X) là nghiệm của phương trình

$$\int_0^{med(X)} f(x)dx = 0,5 \quad \text{hay} \quad 1 - e^{-\frac{[med(X)]^2}{4}} = 0,5$$

Suy ra $med(X) = 1,665$.

⊙ **Chú ý** Nói chung, ba số đặc trưng kỳ vọng, mode và trung vị không trùng nhau. Chẳng hạn, từ các ví dụ (12), (13) và tính thêm kỳ vọng ta có $E(X) = 1,772$; $mod(X) = 1,414$ và $med(X) = 1,665$. Tuy nhiên nếu phân phối đối xứng và chỉ có một mode thì cả ba đặc trưng đó trùng nhau.

2.6 Moment

□ **Định nghĩa 11**

** Moment cấp k của đại lượng ngẫu nhiên X là số $m_k = E(X^k)$.*
** Moment qui tâm cấp k của đại lượng ngẫu nhiên X là số $\alpha_k = E\{[X - E(X)]^k\}$.*

⊕ **Nhận xét**

i) Moment cấp 1 của X là kỳ vọng của X ($m_1 = E(X)$).

ii) Moment qui tâm cấp hai của X là phương sai của X ($\alpha_2 = m_2 - m_1^2 = Var(X)$).

iii) $\alpha_3 = m_3 - 3m_2m_1 + 2m_1^3$.

2.7 Hàm moment sinh

□ **Định nghĩa 12** *Hàm moment sinh của đại lượng ngẫu nhiên X là hàm xác định trong $(-\infty, +\infty)$ cho bởi*

$$\phi(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx}p(x) & \text{nếu } X \text{ rời rạc} \\ \int_{-\infty}^{+\infty} e^{tx}p(x)dx & \text{nếu } X \text{ liên tục} \end{cases}$$

◇ **Tính chất**

i) $\phi'(0) = E(X)$.

ii) $\phi''(0) = E(X^2)$.

iii) Tổng quát: $\phi^{(n)}(0) = E(X^n)$, $\forall n \geq 1$.

Chứng minh.

i) $\phi^{'}(t) = \dfrac{d}{dt}E(e^{tX}) = E\left(\dfrac{d}{dt}(e^{tX})\right) = E(Xe^{tX}).$

Suy ra $\phi^{'}(0) = E(X).$

ii) $\phi^{''}(t) = \dfrac{d}{dt}\phi^{'}(t) = \dfrac{d}{dt}E(Xe^{tX}) = E\left(\dfrac{d}{dt}(Xe^{tX})\right) = E(X^2e^{tX}).$

Suy ra $\phi^{''}(0) = E(X^2).$ □

⊙ **Chú ý**

i) Giả sử X và Y là hai đại lượng ngẫu nhiên độc lập có hàm moment sinh tương ứng là $\phi_X(t)$ và $\phi_Y(t)$. Khi đó hàm moment sinh của $X+Y$ cho bởi

$$\phi_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX}e^{tY}) = E(e^{tX})E(e^{tY}) = \phi_X(t)\phi_Y(t)$$

(đẳng thức gần cuối có được do e^{tX} và e^{tY} độc lập)

ii) Có tương ứng 1−1 giữa hàm moment sinh và hàm phân phối xác suất của đại lượng ngẫu nhiên X.

3. MỘT SỐ QUI LUẬT PHÂN PHỐI XÁC SUẤT

3.1 Phân phối nhị thức (Binomial Distribution)

□ **Định nghĩa 13** *Đại lượng ngẫu nhiên rời rạc X nhận môt trong các giá trị 0,1,2,...,n với các xác suất tương ứng được tính theo công thức Bernoulli*

$$P_x = P(X = x) = C_n^x p^x q^{n-x} \tag{2.1}$$

gọi là có phân phối nhị thức với tham số n và p. Kí hiệu $X \in B(n,p)$ (hay $X \sim B(n,p)$).

⊙ **Công thức**

Với h nguyên dương và $h \leq n - x$, ta có

$$\boxed{P(x \leq X \leq x+h) = P_x + P_{x+1} + \ldots + P_{x+h}} \tag{2.2}$$

• **Ví dụ 14** *Tỷ lệ phế phẩm trong lô sản phẩm là 3%. Lấy ngẫu nhiên 100 sản phẩm để kiểm tra. Tìm xác suất để trong đó*

i) Có 3 phế phẩm.

ii) Có không quá 3 phế phẩm.

Giải

Ta thấy mỗi lần kiểm tra một sản phẩm là thực hiện một phép thử. Do đó ta có n=100 phép thử.

Gọi A là biến cố sản phẩm lấy ra là phế phẩm thì trong mỗi phép thử. Ta có $p = p(A) = 0,03$.

Đặt X là tổng số phế phẩm trong 100 sản phẩm thì $X \in B(100; 0,03)$.

i) $P(X = 3) = C_{100}^3 (0,03)^3.(0,97)^{97} = 0,2274$.

ii)
$$\begin{aligned} P(0 \le X \le 3) &= P_0 + P_1 + P_2 + P_3 \\ &= C_{100}^0 (0,03)^0 (0,97)^{100} + C_{100}^1 (0,03)^1 (0,97)^{99} \\ &\quad + C_{100}^2 (0,03)^2 (0,97)^{98} + C_{100}^3 (0,03)^3 (0,97)^{97} \\ &= 0,647. \end{aligned}$$

⊙ **Chú ý** Khi n khá lớn thì xác suất p không quá gần 0 và 1. Khi đó ta có thể áp dụng công thức xấp xỉ sau

i)
$$\boxed{P_x = C_n^x p^x q^{n-x} \approx \frac{1}{\sqrt{npq}} f(u)} \tag{2.3}$$

trong đó
$$u = \frac{x - np}{\sqrt{npq}}; \quad f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}};$$

(2.3) được gọi công thức địa phương Laplace.

ii)
$$\boxed{P(x \le X \le x + h) \approx \varphi(u_2) - \varphi(u_1)} \tag{2.4}$$

trong đó
$$\varphi(u) = \frac{1}{\sqrt{2\pi}} \int_0^u e^{-\frac{t^2}{2}} dt \quad \text{(Hàm Laplace)};$$
$$u_1 = \frac{x - np}{\sqrt{npq}}; \quad u_2 = \frac{x + h - np}{\sqrt{npq}}$$

(2.4) được gọi là công thức tích phân Laplace.

⊙ **Các tham số đặc trưng**

Nếu $X \in B(n, p)$ thì ta có

i) $E(X) = np$.

ii) $Var(X) = npq$.

iii) $np - q \le mod(X) \le np + p$.

Chứng minh. Xét đại lượng ngẫu nhiên X có phân phối nhị thức với các tham số n và p biểu diễn phép thử biến cố A xảy ra, mỗi phép thử có cùng xác suất xảy ra biến cố A là p.

Ta có thể biểu diễn X như sau:
$$X = \sum_{i=1}^n X_i$$

trong đó $X_i = \begin{cases} 1 & \text{nếu ở phép thử thứ } i \text{ biến cố } A \text{ xảy ra} \\ 0 & \text{nếu ngược lại} \end{cases}$

Vì $X_i,\ i = 1, 2, \ldots, n$ là các đại lượng ngẫu nhiên độc lập có phân phối nhị thức nên

$$E(X_i) = P(X_i = 1) = p$$

$$Var(X_i) = E(X_i^2) - p^2 = p(1-p) = pq \qquad (X_i^2 = X_i)$$

Do đó

$$E(X) = \sum_{i=1}^{n} E(X_i) = np$$

$$Var(X) = \sum_{i=1}^{n} Var(X_i) = npq$$

□

• **Ví dụ 15** *Một máy sản xuất được 200 sản phẩm trong một ngày. Xác suất để máy sản xuất ra phế phẩm là* $0,05$*. Tìm số phế phẩm trung bình và số phế phẩm có khả năng tin chác của máy đó trong một ngày.*

Giải

Gọi X là số phế phẩm của máy trong một ngày thì $X \in B(200; 0,05)$.

Số phế phẩm trung bình của máy trong một ngày là

$$E(X) = np = 200 \times 0,05 = 10$$

Số phế phẩm tin chắc trong ngày là mod(X). Ta có
$np - q = 200 \times 0,05 - 0,95 = 9,05$
$np + p = 200 \times 0,05 + 0,05 = 10,05$

$\Longrightarrow 9,05 \leq mod(X) \leq 10,05$

Vì $X \in B(200; 0,05)$ nên $mod(X) \in Z$. Do đó $mod(X) = 10$.

3.2 Phân phối Poisson

⊙ Công thức Poisson

Giả sử X là đại lượng ngẫu nhiên có phân phối nhị thức với tham số (n, p) và $a = np$ trong đó n khá lớn và p khá bé.

Ta có

$$\begin{aligned} P(X = k) &= \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \\ &= \frac{n!}{(n-k)!k!} . (\frac{a}{n})^k . (1 - \frac{a}{n})^{n-k} \\ &= \frac{n(n-1)\ldots(n-k+1)}{n^k} . \frac{a^k}{k!} . \frac{(1-\frac{a}{n})^n}{(1-\frac{a}{n})^k} \end{aligned}$$

Do n khá lớn và p khá bé nên

$$(1-\frac{a}{n})^n \approx e^{-a}, \qquad \frac{n(n-1)\dots(n-k+1)}{n^k} \approx 1, \qquad (1-\frac{a}{n})^k \approx 1$$

Do đó $P(X=k) \approx e^{-a}\dfrac{a^k}{k!}$

Vậy từ công thức Bernoulli ta có công thức xấp xỉ

$$P_k = P(X=k) = C_n^k p^k q^{n-k} \approx \frac{a^k}{k!}e^{-a}$$

Khi đó ta có thể thay công thức Bernoulli bởi công thức Poisson

$$\boxed{P_k = P(X=k) = \frac{a^k}{k!}e^{-a}} \qquad (2.5)$$

□ **Định nghĩa 14** *Đại lượng ngẫu nhiên rời rạc X nhận một trong các giá trị 0,1,...,n với các xác suất tương ứng được tính theo công thức (2.5) được gọi là có phân phối Poisson với tham số a. Kí hiệu $X \in \mathcal{P}(a)$ (hay $X \sim \mathcal{P}(a)$).*

⊙ **Chú ý**

$$P(k \le X \le k+h) = P_k + P_{k+1} + \dots + P_{k+h} \text{ với } P_k = \frac{a^k}{k!}e^{-a}.$$

• **Ví dụ 16** *Một máy dệt có 1000 ống sợi, Xác suất để một giờ máy hoạt động có 1 ống sợi bị đứt là 0,002. Tìm xác suất để trong một giờ máy hoạt động có không quá 2 ống sợi bị đứt.*

Giải

Việc quan sát một ống sợi có bị đứt hay không trong một giờ máy hoạt động là một phép thử. Máy đệt có 1000 ống sợi nên ta có n = 1000 phép thử độc lập.

Gọi A là biến cố ống sợi bị đứt và X là số ống sợi bị đứt trong một giờ máy hoạt động thì $p = P(A) = 0,002$ và $X \in B(1000; 0,002)$.

Vì $n = 1000$ khá lớn và $np = 2$ không đổi nên ta có thể xem $X \in \mathcal{P}(a)$.

Do đó xác suất để có không quá 2 ống sợi bị đứt trong một giờ là

$P(0 \le X \le 2) = P_0 + P_1 + P_2$

$P_0 = P(X=0) = \frac{2^0}{0!}e^{-2}$

$P_1 = P(X=1) = \frac{2^1}{1!}e^{-2}$

$P_2 = P(X=2) = \frac{2^2}{2!}e^{-2}$

Do đó $P(0 \le X \le 2) = (1+2+2)e^{-2} = 5(2,71)^{-2} = 0,6808$.

⊙ **Các tham số đặc trưng**

Nếu $X \in \mathcal{P}(a)$ thì $E(X) = Var(X) = a$ và $a - 1 \leq modX \leq a$.

Chứng minh. Để nhận được kỳ vọng và phương sai của đại lượng ngẫu nhiên có phân phối Poisson ta xác định hàm moment sinh

$$\psi(t) = E(e^{tX})$$

Ta có

$$\psi(t) = \sum_{k=0}^{\infty} e^{tk} e^{-a} \frac{a^k}{k!} = e^{-a} \sum_{k=0}^{\infty} \frac{(ae^t)^k}{k!} = e^{-a} e^{ae^t} = e^{a(e^t-1)}$$

$\psi^{'}(t) = ae^t e^{a(e^t-1)}$

$\psi^{''}(t) = (ae^t)^2 e^{a(e^t-1)} + ae^t e^{a(e^t-1)}$

Do đó

$$E(X) = \psi^{'}(0) = a$$

$$Var(X) = \psi^{''}(0) - [E(X)]^2 = a^2 + a - a^2 = a$$

□

⊙ **Ửng dụng**

Một vài đại lượng ngẫu nhiên có phân phối Poisson:

i) Số lỗi in sai trong một trang (hoặc một số trang) của một cuốn sách.

ii) Số người trong một cộng đồng sống cho tới 100 tuổi.

iii) Số cuộc điện thoại gọi sai trong một ngày.

iv) Số transitor hư trong ngày đầu tiên sử dụng.

v) Số khách hàng vào bưu điện trong một ngày.

vi) Số hạt α phát ra từ cát hạt phóng xạ trong một chu kỳ.

3.3 Phân phối siêu bội

a) Công thức siêu bội

Xét một tập hợp gồm N phần tử, trong đó có M phần tử có tính chất A nào đó. Lấy ngẫu nhiên (không hoàn lại) từ tập hợp ra n phần tử. Gọi X là số phần tử có tính chất A có trong n phần tử lấy ra. Ta có

$$\boxed{P_x = P(X = x) = \frac{C_M^x C_{N-M}^{n-x}}{C_N^n}} \quad (x = 0, 1, \ldots, n) \tag{2.6}$$

b) Phân phối siêu bội

□ **Định nghĩa 15** *Đại lượng ngẫu nhiên rời rạc X nhận một trong các giá trị 0,1,...,n với các xác suất tương ứng được tính theo công thức (2.6) được gọi là có phân phối siêu bội với tham số N, M, n. Kí hiệu $X \in H(N,M,n)$ (hay $X \sim H(N,M,n)$).*

• **Ví dụ 17** *Một lô hàng có 10 sản phẩm, trong đó có 6 sản phẩm tốt. Lấy ngẫu nhiên (không hoàn lại) từ lô hàng ra 4 sản phẩm. Tìm xác suất để có 3 sản phẩm tốt trong 4 sản phẩm được lấy ra.*

Giải

Gọi X là số sản phẩm tốt có trong 4 sản phẩm lấy ra thì X là đại lượng ngẫu nhiên có phân phối siêu bội với tham số $N = 10, M = 6, n = 4$.

Xác suất để có 3 sản phẩm tốt trong 4 sản phẩm lấy ra là

$$P(X = 3) = \frac{C_6^3.C_4^1}{C_{10}^4} = \frac{8}{21} = 0,3809$$

⊙ **Chú ý**

Khi n khá bé so với N thì $\dfrac{C_M^x C_{N-M}^{n-x}}{C_N^n} \approx C_n^x p^x q^{n-x} \quad (p = \dfrac{M}{N},\ q = 1 - p)$

Gọi X là số phần tử có tính chất A nào đó trong n phần tử lấy ra thì ta có thể xem $X \in B(n,p)$ vói p là tỉ lệ phần tử có tính chất A của tập hợp.

c) Các tham số đặc trưng

Nếu $X \in H(N,M,n)$ thì ta có

$$E(X) = np \quad (\text{với } p = \frac{M}{N})$$

$$Var(X) = npq\frac{N-n}{N-1} \quad (\text{với } q = 1 - p).$$

Bảng tổng kết các phân phối rời rạc

Phân phối	Kí hiệu	Xác suất $P(X = k)$	$E(X)$	$Var(X)$
Nhị thức	$B(n,p)$	$C_n^k p^k (1-p)^{n-k}$	np	npq
Poisson	$\mathcal{P}(a)$	$\dfrac{a^k}{k!}e^{-a}$	a	a
Siêu bội	$H(N,M,n)$	$\dfrac{C_M^k.C_{N-M}^{n-k}}{C_N^n}$	$np\ (p = \frac{M}{N})$	$npq\dfrac{N-n}{N-1}$

3.4 Phân phối mũ

□ **Định nghĩa 16** *Đại lượng ngẫu nhiên X được gọi là có phân phối mũ với tham số $\lambda > 0$ nếu nó có hàm mật độ xác suất*

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{nếu } x > 0 \\ 0 & \text{nếu } x \leq 0 \end{cases}$$

⊕ **Nhận xét** Nếu X có phân phối mũ với tham số λ thì hàm phân phối xác suất của X là

$F(x) = \int\limits_0^x \lambda e^{-\lambda x} dt = 1 - e^{-\lambda x}$ vói $x > 0$

và

$F(x) = 0$ với $x \leq 0$.

⊙ **Các tham số đặc trưng**

Nếu X là đại lượng ngẫu nhiên có phân phối mũ với tham số $\lambda > 0$ thì

i) Kỳ vọng của X là

$$E(X) = \lambda \int_0^{+\infty} x e^{-\lambda x} dx = \left[-x e^{-\lambda x}\right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx = \frac{1}{\lambda}$$

ii) Phương sai của X là

$$Var(X) = \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2}$$

Tích phân từng phần ta được $\int\limits_0^{+\infty} x^2 \lambda e^{-\lambda x} dx = \left[-x^2 e^{-\lambda x}\right]_0^{+\infty} + 2 \int\limits_0^{+\infty} \lambda x e^{-\lambda x} dx = \frac{2}{\lambda^2}$.

Do đó $Var(X) = \frac{1}{\lambda^2}$.

• **Ví dụ 18** *Giả sử tuổi thọ (tính bằng năm) của một mạch điện tử trong máy tính là một đại lượng ngẫu nhiên có phân phối mũ với kỳ vọng là 6,25. 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. Thì X có phân phối mũ

Ta có $\lambda = \frac{1}{E(X)} = \frac{1}{6,25}$

$P(X \leq 5) = F(5) = 1 - e^{-\lambda .5} = 1 - e^{-\frac{5}{6,25}} = 1 - e^{-0,8} = 1 - 0,449 = 0,5506$

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

⊙ **Ứng dụng trong thực tế**

Khoảng thời gian giữa hai lần xuất hiện của một biến có phân phối mũ. Chẳng hạn khoảng thời gian giữa hai ca cấp cứu ở một bệnh viện, giữa hai lần hỏng hóc của một cái máy, giữa hai trận lụt hay động đất là những đại lượng ngẫu nhiên có phân phối mũ.

3.5 Phân phối đều

□ **Định nghĩa 17** *Đại lượng ngẫu nhiên liên tục X được gọi là có phân phối đều trên đoạn [a,b] nếu hàm mật độ xác suất có dạng*

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{nếu} \quad x \in [a,b] \\ 0 & \text{nếu} \quad x \notin [a,b] \end{cases}$$

⊕ **Nhận xét** Nếu X có phân phối đều trên [a,b] thì hàm phân phối của X cho bởi

$F(x) = 0$ nếu $x < a$

$F(x) = \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{a}^{x} \dfrac{dx}{b-a} = \dfrac{x-a}{b-a}$ nếu $a \le x \le b$

$F(x) = 1$ nếu $x > b$.

⊙ **Chú ý** Giả sử $(\alpha, \beta) \subset [a, b]$. Xác suất để X rơi vào (α, β) là

$$P(\alpha < X < \beta) = \int\limits_{\alpha}^{\beta} f(x)dx = \frac{\beta - \alpha}{b-a}$$

⊙ **Các tham số đăc trưng**

i) $E(X) = \int\limits_{a}^{b} \dfrac{xdx}{b-a} = \dfrac{1}{b-a}\dfrac{b^2-a^2}{2} = \dfrac{a+b}{2}$ (kỳ vọng là trung điểm của [a,b]).

ii) $$\begin{aligned} Var(X) &= \int\limits_{a}^{b} \frac{x^2dx}{b-a} - [E(X)]^2 = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b - \left(\frac{a+b}{2}\right) \\ &= \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12} \end{aligned}$$

iii) modX là bất cứ điểm nào trên [a,b].

• **Ví dụ 19** *Lịch chạy của xe buýt tại một trạm xe buýt như sau: chiếc xe buýt đầu tiên trong ngày sẽ khởi hành từ trạm này vào lúc 7 giờ, cứ sau mỗi 15 phút sẽ có một xe khác đến trạm. Giả sử một hành khách đến trạm 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.

Giải

Gọi X là số phút sau 7 giờ mà hành khách đến trạm thì X là đại lượng ngẫu nhiên có phân phối đều trong khoảng $(0, 30)$.

a) Hành khách sẽ 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

$$P(10 < X < 15) + P(25 < X < 30) = \frac{5}{30} + \frac{5}{30} = \frac{1}{3}$$

b) Hành khách chờ ít nhất 12 phút nếu đến trạm giữa 7giờ và 7 giờ 3 phút hoặc giữa 7 giờ 15 phút và 7 giờ 18 phút. Xác suất cần tìm là

$$P(0 < X < 3) + P(15 < X < 18) = \frac{3}{30} + \frac{3}{30} = \frac{1}{5}$$

3.6 Phân phối chuẩn (Karl Gauss)

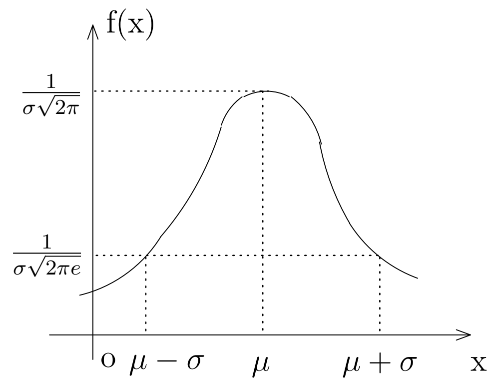
a) Phân phối chuẩn

□ **Định nghĩa 18**

Đại lượng ngẫu nhiên liên tục X nhận giá trị trong khoảng $(-\infty, +\infty)$ được gọi là có phân phối chuẩn nếu hàm mật độ xác suất có dạng

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

trong đó μ, σ là hằng số, $\sigma > 0$, $-\infty < x < \infty$.



Kí hiệu $X \in N(\mu, \sigma^2)$ hay $(X \sim N(\mu, \sigma^2))$.

b) Các tham số đặc trưng

Nếu $X \in N(\mu, \sigma^2)$ thì $E(X) = \mu$ và $Var(X) = \sigma^2$.

Chứng minh. Xét hàm moment sinh

$$\phi(t) = E(e^{tX}) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tx}.e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Đặt $y = \frac{x-\mu}{\sigma}$ thì

$$\begin{aligned}\phi(t) &= \frac{1}{\sqrt{2\pi}}e^{\mu t}\int\limits_{-\infty}^{+\infty} e^{tx}e^{-\frac{y^2}{2}}dy &&= \frac{e^{\mu t}}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{y^2-2t\sigma y}{2}}dy \\ &= \frac{e^{\mu t}}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{(y-t\sigma)^2}{2}+\frac{t^2\sigma^2}{2}}dy &&= e^{\mu t+\frac{\sigma^2t^2}{2}} \times \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{(y-t\sigma)^2}{2}}dy\end{aligned}$$

Vì $f(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(y-t\sigma)^2}{2}}$ là hàm mật độ của phân phối chuẩn với tham số $t\sigma$ và 1 nên $\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{(y-t\sigma)^2}{2}}dy = 1$.

Do đó $\phi(t) = e^{\mu t+\frac{\sigma^2+t^2}{2}}$.

Lấy các đạo hàm ta được

$$\phi^{'}(t) = (\mu + t\sigma^2)e^{\mu t+\sigma^2\frac{t^2}{2}}, \quad \phi^{''}(t) = \sigma^2 e^{\mu t+\sigma^2\frac{t^2}{2}}.(\mu + t\sigma^2)$$

Khi đó

$E(X) = \phi^{'}(0) = \mu$

$E(X^2) = \phi^{''}(0) = \sigma^2 + \mu^2 \quad \Longrightarrow \quad Var(X) = E(X^2) - [E(X)]^2 = \sigma^2$ □

c) Phân phối chuẩn hóa

□ **Định nghĩa 19** *Đại lượng ngẫu nhiên X được gọi là có phân phối chuẩn hóa nếu nó có phân phối chuẩn với $\mu = 0$ và $\sigma^2 = 1$. Kí hiệu $X \in N(0,1)$ hay $X \sim N(0,1)$.*

⊕ **Nhận xét** Nếu $X \in N(\mu, \sigma^2)$ thì $U = \dfrac{X-\mu}{\sigma} \in N(0,1)$.

d) Phân vị chuẩn

Phân vị chuẩn mức α, kí hiệu u_α, là giá trị của đại lượng ngẫu nhiên U có phân phối chuẩn hóa thỏa mãn điều kiện

$$P(U < u_\alpha) = \alpha.$$

Với α cho trước có thể tính được các giá trị của u_α. Các giá trị của u_α được tính sẵn thành bảng.

e) Công thức

Nếu $X \in N(\mu, \sigma^2)$ thì ta có

i) $P(x_1 \leq X \leq x_2) = \varphi(\frac{x_2 - \mu}{\sigma}) - \varphi(\frac{x_1 - \mu}{\sigma})$

ii) $P(|X - \mu| < \varepsilon) = 2\varphi(\frac{\varepsilon}{\sigma})$

trong đó $\varphi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt$ (hàm Laplace).

• **Ví dụ 20** *Trọng lượng của một loại sản phẩm là đại lượng ngẫu nhiên có phân phối chuẩn với trọng lượng trung bình $\mu = 5kg$ và đọ lệch tiêu chuẩn $\sigma = 0,1$. Tính tỉ lệ những sản phẩm có trọng lượng từ 4,9 kg đến 5,2 kg.*

Giải

Gọi X là trọng lượng của sản phẩm thì $X \in N(5; 0,1)$.

Tỉ lệ sản phẩm có trọng lượng từ 4,9 kg đến 5,2 kg là

$$\begin{aligned} P(4,9 \leq X \leq 5,2) &= \varphi(\tfrac{5,2-5}{0,1}) - \varphi(\tfrac{4,9-5}{0,1}) \\ &= \varphi(2) - \varphi(-1) \\ &= 0,4772 - (-0,3413) \\ &= 0,8185 \end{aligned}$$

f) Qui tăc "k$-\sigma$"

Trong công thức $P(|X - \mu| < \varepsilon) = 2\varphi(\frac{\varepsilon}{\sigma})$ nếu lấy $\varepsilon = k\sigma$ thì $P(|X - \mu| < \varepsilon) = 2\varphi(k)$.

Trong thực tế ta thường dùng qui tắc $1,96\sigma$, $2,58\sigma$ và 3σ với nội dung là:

"Nếu $X \in N(\mu, \sigma^2)$ thì xác suất để X nhận giá trị sai lệch so với kỳ vọng không quá $1,96\sigma; 2,58\sigma$ và 3σ là 95 %, 99% và 99% ".

g) Ứng dụng

Các đại lượng ngẫu nhiên sau có phân phối chuẩn:

- Kích thước chi tiết máy do máy sản suất ra.
- Trọng lượng của nhều sản phẩm cùng loại.
- Năng suất của một loại cây trồng trên những thửa ruộng khác nhau.

3.7 Phân phối χ^2

□ **Định nghĩa 20** *Giả sử X_i (i=1,2,...,n) là các đại lượng ngẫu nhiên độc lập cùng có phân phối chuẩn hóa.*

Đại lượng ngẫu nhiên $\chi^2 = \sum_{i=1}^{n} X_i^2$ *được gọi là có phân phối* χ^2 *(khi−bình phương) với n bậc tự do. Kí hiệu* $\chi^2 \in \chi^2(n)$ *(hay* $\chi^2 \sim \chi^2(n)$*).*

⊕ **Nhận xét**

Hàm mật độ xác suất của χ^2 có dạng

$$f_n(x) = \begin{cases} \dfrac{e^{-\frac{x}{2}}.x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}.\Gamma(\frac{n}{2})} & \text{với} \quad x > 0 \\ 0 & \text{với} \quad x \le 0 \end{cases}$$

trong đó $\Gamma(x) = \int\limits_0^{+\infty} t^{x-1}e^{-t}dt$ (Hàm Gamma)

Hàm mật độ xác suất của χ^2 với n bậc tự do

⊙ **Các tham số đặc trưng**

Nếu $\chi^2 \in \chi^2(n)$ thì $E(\chi^2) = n$ và $Var(\chi^2) = 2n$.

⊙ **Phân vị** χ^2

Phân vị χ^2 mức α, kí hiệu χ^2_α, là giá trị của đại lượng χ^2_α có phân phối "khi−bình phương" với n bậc tự do thỏa mãn

$$P(\chi^2 < \chi^2_\alpha) = \alpha$$

Các giá trị của χ^2_α được tính sẵn thành bảng.

⊙ **Chú ý** Khi bậc n tăng lên thì phân phối χ^2 xấp xỉ với phân phối chuẩn.

3.8 Phân phối Student (G.S Gosset)

□ **Định nghĩa 21** *Giả sử U là đại lượng ngẫu nhiên có phân phối chuẩn hóa và V là đại lượng ngẫu nhiên độc lập với U có phân phối* χ^2 *với n bậc tự do. Khi đó đại lượng ngẫu nhiên*

$$T = \frac{U\sqrt{n}}{\sqrt{V}}$$

được gọi là có phân phối Student với n bậc tự do. Kí hiệu $T \in T(n)$ *(hay* $T \sim T(n)$*).*

⊕ **Nhận xét** Hàm mật độ của đại lượng ngẫu nhiên có phân phối Student với n bậc tự do có dạng

$$f_n(t) = \frac{\Gamma(\frac{n+1}{2})(1 + \frac{t^2}{n})^{-\frac{n+1}{2}}}{\Gamma(\frac{n}{2})\sqrt{n\pi}}; \quad (-\infty < t < +\infty)$$

trong đó $\Gamma(x) = \int\limits_0^{+\infty} t^{x-1}e^{-t}dt$ (Hàm Gamma)

⊙ **Các tham số đặc trưng**

Nếu $T \in T(n)$ thì $E(T) = 0$ và $Var(T) = \dfrac{n}{n-2}$.

• **Phân vị Student**

Phân vị Student mức α, kí hiệu t_α là giá trị của đại lượng ngẫu nhiên $T \in T(n)$ thỏa mãn $P(T < t_\alpha) = \alpha$.

Ta có $t_\alpha = -t_{1-\alpha}$.

⊙ **Chú ý**

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. Các biến có về giá và thời gian thường giới hạn một cách nghiêm ngặt kích thước của mẫu. Chính vì thế phân phối chuẩn không thể dùng để xấp xỉ phân phối khi mẫu có kích thước nhỏ. Trong trường hợp này ta dùng phân phối Student.

Khi bậc tự do n tăng lên ($n > 30$) thì phân phối Student tiến nhanh về phân phối chuẩn. Do đó khi $n > 30$ ta có thể dùng phân phối chuẩn thay cho phân phối Student.

3.9 Phân phối F (Fisher−Snedecor)

□ **Định nghĩa 22** *Nếu χ^2_n và χ^2_m là hai đại lượng ngẫu nhiên có phân phối "khi bình phương" với n và m bậc tự do thì đại lượng ngẫu nhiên $F_{n,m}$ xác định bởi*

$$F_{n,m} = \frac{\chi^2_n/n}{\chi^2_m/m}$$

được gọi là có phân phối F với n và m bậc tự do.

⊕ **Nhận xét** Hàm mật độ của phân phối F có dạng

$$p(x) = \begin{cases} 0 & ; \quad x \le 0 \\ \frac{\Gamma(\frac{n+m}{2})}{\Gamma(\frac{n}{2}).\Gamma(\frac{m}{2})}\left(\frac{n}{m}\right)^{\frac{n}{2}} \frac{x^{\frac{n}{2}-1}}{(1+\frac{n}{m}x)^{\frac{n+m}{2}}} & ; \quad x > 0 \end{cases}$$

• **Các tham số đặc trưng**

$E(F_{n,m}) = \dfrac{m}{m-2}$ với $m > 2$

$Var(F_{n,m}) = \dfrac{m^2(2m+2n-4)}{n(m-2)^2(m-4)}$ với $m > 4$

3.10 Phân phối Gamma

□ **Định nghĩa 23** *Đại lượng ngẫu nhiên X được gọi là có phân phối Gamma với các tham số (α, λ), kí hiệu $X \in \gamma(\alpha, \lambda)$, nếu hàm mật độ xác suất có dạng*

$$f(x) = \begin{cases} \dfrac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & ; \quad x \geq 0 \\ 0 & ; \quad x < 0 \end{cases}$$

trong đó $\Gamma(\alpha) = \int\limits_0^\infty \lambda e^{-\lambda x}(\lambda x)^{\alpha-1}dx = \int\limits_0^\infty e^{-y}y^{\alpha-1}dy \quad (y = \lambda x)$.

⊙ **Các tham số đặc trưng**

Nếu $X \in \gamma(\alpha, \lambda)$ thì $E(X) = \dfrac{\alpha}{\lambda}$ và $Var(X) = \dfrac{\alpha}{\lambda^2}$.

◇ **Tính chất** Nếu $X \in \gamma(\alpha, \lambda)$ và $Y \in \gamma(\beta, \lambda)$ thì $X + Y \in \gamma(\alpha + \beta, \lambda)$.

Bảng tổng kết các phân phối liên tục

Phân phối	Kí hiệu	Hàm mật độ $f(x)$	$E(X)$	$Var(X)$
Mũ		$\lambda e^{-\lambda x}$ $(x > 0)$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
Đều		$\dfrac{1}{b-a}$ $(a \leq x \leq b)$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Chuẩn	$N(\sigma^2, \mu)$	$\dfrac{1}{\sigma\sqrt{2\pi}}exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$	μ	σ^2
Khi bình phương	$\chi^2(n)$	$\dfrac{e^{-\frac{x}{2}}.x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}.\Gamma(\frac{n}{2})}$ $(x > 0, n > 0$	n	2n
Student	$T(n)$	$\dfrac{\Gamma(\frac{n+1}{2})(1+\frac{x^2}{n})^{-\frac{n+1}{2}}}{\Gamma(\frac{n}{2})\sqrt{n\pi}}$ $(n > 0)$	0 $(n > 1)$	$\dfrac{n}{n-2}$
Gamma	$\gamma(\alpha, \lambda)$	$\dfrac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}$	$\dfrac{\alpha}{\lambda}$	$\dfrac{\alpha}{\lambda^2}$

4. ĐẠI LƯỢNG NGẪU NHIÊN HAI CHIỀU

4.1 Khái niệm về đại lượng ngẫu nhiên hai chiều

Đại lượng ngẫu nhiên hai chiều là đại lượng ngẫu nhiên mà các giá trị có thể của nó được xác định bằng hai số. Kí hiệu (X, Y).

(X, Y được gọi là các thành phần của đại lượng ngẫu nhiên hai chiều)

Đại lượng ngẫu nhiên hai chiều được gọi là rời rạc (liên tục) nếu các thành phần của nó là các đại lượng ngẫu nhiên rời rạc (liên tục).

4.2 Phân phối xác suất của đại lượng ngẫu nhiên hai chiều

a) Bảng phân phối xác suất

X\Y	y_1	y_2	...	y_j	...	y_m
x_1	$P(x_1, y_1)$	$P(x_2, y_2)$	...	$P(x_1, y_j)$	...	$P(x_1, y_m)$
x_2	$P(x_2, y_1)$	$P(x_2, y_2)$	...	$P(x_2, y_j)$	...	$P(x_2, y_m)$
$\vdots$	...	...	...	...	...	...
x_i	$P(x_i, y_1)$	$P(x_i, y_2)$	...	$P(x_i, y_j)$	...	$P(x_i, y_m$
$\vdots$	...	...	...	...	...	...
x_n	$P(x_n, y_1)$	$P(x_n, y_2)$	...	$P(x_n, y_j)$	...	$P(x_n, y_m)$

trong đó

$x_i\,(i = \overline{1, n})$ là các giá trị có thể của thành phần X

$y_j\,(j = \overline{1, m})$ là các giá trị có thể của thành phần Y

$P(x_i, y_j) = P(\,(X, Y) = (x_i, y_j)\,) = P(X = x_i, Y = y_j), \quad i = \overline{1, n}, j = \overline{1, m}$

$\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} P(x_i, y_j) = 1$

b) Hàm mật độ xác suất

□ **Định nghĩa 24** *Hàm không âm, liên tục $f(x, y)$ được gọi là hàm mật độ xác suất của đại lượng ngẫu nhiên hai chiều (X, Y) nếu nó thỏa mãn*

$$P(X \in A, Y \in B) = \int\limits_A dx \int\limits_B f(x, y)dy$$

với A, B là các tập số thực.

c) Hàm phân phối xác suất

□ **Định nghĩa 25** *Hàm phân phối xác suất của đại lượng ngẫu nhiên hai chiều (X, Y), kí hiệu $F(x, y)$, là hàm được xác định như sau*

$$F(x, y) = P(X < x, Y < y)$$

⊙ **Nhận xét**

Ta có $F(x, y) = P(X < x, Y < y) = \int\limits_{-\infty}^{x}\left(\int\limits_{-\infty}^{y} f(x, y)dy\right)dx$ nên

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$$

4.3 Kỳ vọng và phương sai của các thành phần

i) Trường hợp (X, Y) rời rạc

$$E(X) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i P(x_i, y_j); \quad E(Y) = \sum_{j=1}^{m}\sum_{i=1}^{n} y_j P(x_i, y_j)$$

$$Var(X) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^2 P(x_i, y_j) - [E(X)]^2, \quad Var(Y) = \sum_{j=1}^{m}\sum_{i=1}^{n} y_j^2 P(x_i, y_j) - [E(Y)]^2$$

ii) Trường hợp (X, Y) liên tục

$$E(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f(x, y)dxdy, \quad E(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f(x, y)dxdy.$$

$$Var(X) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2 f(x, y)dxdy - [E(X)]^2, \quad Var(Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2 f(x, y)dxdy - [E(Y)]^2$$

5. PHÂN PHỐI XÁC SUẤT CỦA HÀM CÁC ĐẠI LƯỢNG NGẪU NHIÊN

5.1 Hàm của một đại lượng ngẫu nhiên

□ **Định nghĩa 26** *Nếu mỗi giá trị có thể của đại lượng ngẫu nhiên X tương ứng với một giá trị có thể của đại lượng ngẫu nhiên Y thì Y được gọi là hàm của đại lượng ngẫu nhiên X. Kí hiệu* $Y = \varphi(X)$.

◇ **Tính chất**

i) Nếu X là đại lượng ngẫu nhiên rời rạc và $Y = \varphi(X)$ thì ứng với các giá trị khác nhau của X ta có các giá trị khác nhau của Y và có

$$P(Y = \varphi(x_i)) = P(X = x_i)$$

ii) Giả sử X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$ và $Y = \varphi(X)$.

Nếu $y = \varphi(x)$ là hàm khả vi, đơn điệu, có hàm ngược là $x = \psi(y)$ thì hàm mật độ xác suất $g(y)$ của đại lượng ngẫu nhiên Y được xác định bởi

$$g(y) = f(\psi(y)).\psi'(y)$$

• **Ví dụ 21** *Giả sử X là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất*

X	1	3	4
P	0,3	0,5	0,2

Tìm qui luật phân phối xác suất của $Y = X^2$.

Giải

Các giá trị Y có thể nhận là $y_1 = 1^2 = 1$; $y_2 = 3^2 = 9$; $y_3 = 4^2 = 16$. Vậy phân phối xác suất của Y có thể cho bởi

Y	1	9	16
P	0,3	0,5	0,2

• **Các tham số**

i) Nếu X là đại lượng ngẫu nhiên rời rạc nhận một trong các giá trị $x_1, x_2, \ldots, x_n$ với các xác suất tương ứng $p_1, p_2, \ldots, p_n$ thì

$$E(Y) = E[\varphi(X)] = \sum_{i=1}^{n} \varphi(x_i)p_i$$

$$Var(Y) = Var[\varphi(X)] = \sum_{i=1}^{n} \varphi^2(x_i)p_i - [E(Y)]^2$$

ii) Nếu X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$ thì

$$E(Y) = E[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x)f(x)dx$$

$$Var(Y) = Var[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi^2(x)f(x)dx - [E(Y)]^2$$

5.2 Hàm của đại lượng ngẫu nhiên hai chiều

□ **Định nghĩa 27** *Nếu mỗi cặp giá trị có thể các đại lượng X và Y tương ứng với một giá trị có thể của Z thì Z được gọi là hàm của hai đại lượng ngẫu nhiên X, Y. Kí hiệu* $Z = \varphi(X, Y)$.

⊙ **Chú ý** Việc xác định phân phối xác suất của $Z = \varphi(X, Y)$ thường rất phức tạp. Ta xét trường hợp đơn giản $Z = X + Y$ thông qua ví dụ dưới đây.

• **Ví dụ 22** *Giả sử X và Y là hai đại lượng ngẫu nhiên độc lập có bảng phân phối xác suất*

X	1	2
P	0,3	0,7

Y	3	4
P	0,2	0,8

Tìm phân phối xác suất của $Z = X + Y$.

Giải

Các giá trị có thể của Z là tổng của một giá trị của X và một giá trị có thể của Y.

Do đó Z nhận các giá trị có thể

$z_1 = 1 + 3 = 4$; $z_2 = 1 + 4 = 5$; $z_3 = 2 + 3 = 5$; $z_4 = 2 + 4 = 6$

Các xác suất tương ứng là

$P(Z = 4) = P(X = 1).P(Y = 3) = 0,3 \times 0,2 = 0,06$

$P(Z = 5) = P(X = 1, Y = 4) + P(X = 2, Y = 3)$

$= P(X = 1).P(Y = 4) + P(X = 2).P(Y = 3)$

$= 0,3 \times 0,8 + 0,7 \times 0,2 = 0,38$

$P(Z = 6) = P(X = 2).P(Y = 4) = 0.7 \times 0,8 = 0,56$

Vậy Z có phân phối xác suất

Z	4	5	6
P	0,006	0,38	0,56

6. LUẬT SỐ LỚN

6.1 Bất đẳng thức Markov

△ **Định lý 1** *Nếu X là đại lượng ngẫu nhiên nhận giá trị không âm thì* $\forall \varepsilon > 0$ *ta có*

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Chứng minh. Ta chứng minh trong trường hợp X là đại lượng ngẫu nhiên liên tục có hàm mật độ $f(x)$.

$$E(X) = \int_0^{+\infty} xf(x)dx = \int_0^{a} xf(x)dx + \int_a^{+\infty} xf(x)dx$$

$$\geq \int_a^{+\infty} xf(x)dx \geq \int_a^{+\infty} af(x)dx = a\int_a^{+\infty} = aP(X \geq a).$$

□

6.2 Bất đẳng thức Tchebyshev

△ **Định lý 2** *Nếu X là đại lượng ngẫu nhiên có kỳ vọng* μ *và phương sai* σ^2 *hữu hạn thì* $\forall \varepsilon > 0$ *bé tùy ý ta có*

$$P(|X - \mu| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

hay

$$P(|X - \mu| < \varepsilon) > 1 - \frac{Var(X)}{\varepsilon^2}$$

Chứng minh.

Ta thấy $(X - \mu)^2$ là đại lượng ngẫu nhiên nhận giá trị không âm.

Áp dụng bất đẳng thức Tchebyshev với $a = \varepsilon^2$ ta được

$$P[(X - \mu)^2 \geq \varepsilon^2] \leq \frac{E[(X - \mu)^2]}{\varepsilon^2} = \frac{Var(X)}{\varepsilon^2}$$

Vì $(X-\mu)^2 \geq \varepsilon^2$ khi và chỉ khi $|X-\mu| \geq \varepsilon$ nên

$$P(|X-\mu| \geq \varepsilon) \geq \frac{Var(X)}{\varepsilon^2}$$

□

⊙ **Chú ý** Bất đẳng thức Markov và Tchebuchev giúp ta phương tiện thấy được giới hạn của xác suất khi kỳ vọng và phương sai của phân phối xác suất chưa biết.

• **Ví dụ 23** *Giả sử số sản phẩm được sản xuất của một nhà máy trong một tuần là một đại lượng ngẫu nhiên với kỳ vọng $\mu = 50$.*

a) Có thể nói gì về xác suất sản phẩm của tuần này vượt quá 75.

b) Nếu phương sai của sản phẩm trong tuần này là $\sigma^2 = 25$ thì có thể nói gì về xác suất sản phẩm tuần này sẽ ở giữa 40 và 60.

Giải

a) Theo bất đẳng thức Markov

$$P(X > 75) \geq \frac{E(X)}{75} = \frac{50}{75} = \frac{2}{3}$$

b) Theo bất đẳng thức Tchebyshev

$$P(|X-50| \geq 10) \leq \frac{\sigma^2}{10^2} = \frac{25}{100} = \frac{1}{4}$$

Do đó

$$P(40 < X < 60) = P(|X-50| < 10) > 1 - \frac{1}{4} = \frac{3}{4}$$

6.3 Định lý Tchebyshev

∆ **Định lý 3** (`Định lý Tchebyshev`) *Nếu các đại lượng ngẫu nhiên $X_1, X_2, \ldots, X_n$ độc lập từng đôi, có kỳ vọng hữu hạn và các phương sai đều bị chặn trên bởi số C thì $\forall \varepsilon > 0$ bétùy ý ta có*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} E(X_i)\right| < \varepsilon)\right) = 1$$

Đặc biệt, khi $E(X_i) = a;\ (i = \overline{1,n})$ *thì* $\lim\limits_{n\to\infty}(|\frac{1}{n}\sum_{i=1}^{n} X_i - a| < \varepsilon) = 1$

Chứng minh. Ta chứng minh trong trường hợp đặc biệt $E(X_i) = \mu,\ Var(X_i) = \sigma^2$ $(i = 1, 2\ldots, n)$. Ta có

$$E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \mu, \quad Var(\frac{1}{n}\sum_{i=1}^{n}) = \frac{\sigma^2}{n}$$

Theo bất đẳng thức Tchebyshev

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right|\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

□

• **Ý nghĩa**

Mặc dù từng đại lượng ngẫu nhiên độc lập có thể nhận giá trị sai khác nhiều so với kỳ vọng của chúng, nhưng trung bình số học của một số lớn đại lượng ngẫu nhiên lại nhận giá trị gần bằng trung bình số học của các kỳ vọng của chúng. Điều này cho phép ta dự đoán giá trị trung bình số học của các đại lượng ngẫu nhiên.

6.4 Định lý Bernoulli

△ **Định lý 4 (Định lý Bernoulli)** *Nếu f_n là tần suất xuất hiện biến cố A trong n phép thử độc lập và p là xác suất xuất hiện biến cố A trong mỗi phép thử thì $\forall \varepsilon > 0$ bé tùy ý ta có*

$$\lim_{n\to\infty} P(|f_n - p| < \varepsilon) = 1$$

• **Ý nghĩa**

Tần suất xuất hiện biến cố trong n phép thử độc lập dần về xác suất xuất hiện biến cố trong mỗi phép thử khi số phép thử tăng lên vô hạn.

7. BÀI TẬP

1. Một nhóm có 10 người gồm 6 nam và 4 nữ. Chọn ngẫu nhiên ra 3 người. Gọi X là số nữ ở trong nhóm. Lập bảng phân phối xác suất của X và tính $E(X), Var(X), mod(X)$.

2. Gieo đồng thời hai con xúc sắc cân đối đồng chất. Gọi X là tổng số nốt xuất hiện trên hai mặt con xúc sắc. lập bảng qui luật phân phối xác suất của X. Tính $E(X)$ và $Var(X)$.

3. Trong một cái hộp có 5 bóng đèn trong đó có 2 bóng tốt và 3 bóng hỏng. Chọn ngẫu nhiên từng bóng đem thử (thử xong không trả lại) cho đến khi thu được 2 bóng tốt. Gọi X là số lần thử cần thiết. Tìm phân phối xác suất của X. Trung bình cần thử bao nhiêu lần?

4. Một đợt xổ số phát hành N vé. Trong đó có m_i vé trúng k_i đồng một vé ($i = 1, 2, \ldots, n$). Hỏi giá của mỗi vé số là bao nhiêu để cho trung bình của tiền thưởng cho mỗi vé bằng một nửa giá tiền của một vé?

5. Tuổi thọ của một loại côn trùng nào đó là một đại lượng ngẫu nhiên liên tục X (đơn vị là tháng) có hàm mật độ

$$f(x) = \begin{cases} kx^2(4-x) & \text{nếu } 0 \le x \le 4 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a) Tìm hằng số k.
b) Tìm $mod(X)$.
c) Tính xác suất để côn trùng chết trước khi nó được 1 tháng tuổi.

6. Cho đại lượng ngẫu nhiên liên tục X có hàm mật độ

$$f(x) = \begin{cases} kx^2e^{-2x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

a) Tìm hằng số k.
b) Tìm hàm phân phối của X.
c) Tìm $mod(X)$.
d) Tìm $E(X)$ và $Var(X)$.

7. Một xí nghiệp sản xuất máy tính có xác suất làm ra phế phẩm là 0,02. Chọn ngẫu nhiên 250 máy tính để kiểm tra. Tìm xác suất để:
a) Có đúng 2 phế phẩm.
b) Có không quá 2 phế phẩm.

8. `(Bài toán Samuel−Pepys)` Pepys đã đưa ra bài toán sau cho Newton: *Biến cố nào trong các biến có sau đây có xác suất lớn nhất?*

a) Có ít nhất một lần xuất hiện mặt 6 khi tung một con xúc xắc 6 lần.
b) Có ít nhất 2 lần xuất hiện mặt 6 khi tung con xúc xắc 12 lần.
c) Có ít nhất 3 lần xuất hiện mặt 6 khi tung con xúc xắc 18 lần.

9. Xác suất một người bị phản ứng từ việc tiêm huyết thanh là 0,001. Tìm xác suất sao cho trong 2000 người có đúng 3 người, có nhiều hơn 2 người bị phản ứng.

10. Một lô hàng có 500 sản phẩm (trong đó có 400 sản phẩm loại A). Lấy ngẫu nhiên từ lô hàng đó ra 200 sản phẩm để kiểm tra. Gọi X là số sản phẩm loại A có trong 200 sản phẩm lấy ra kiểm tra. Tìm kỳ vọng và phương sai của X.

11. Một trung tâm bưu điện nhận được trung bình 300 lần gọi điện thoại trong một giờ. Tìm xác suất để trung tâm này nhận được đúng 2 lần gọi trong 1 phút.

12. Trọng lượng của một con bò là một đại lượng ngẫu nhiên có phân phối chuẩn với giá trị trung bình $250kg$ và độ lệch tiêu chuẩn là $40kg$. Tìm xác suất để một con bò chọn ngẫu nhiên có trọng lượng:

a) Nặng hơn $300kg$.
b) Nhẹ hơn $175kg$.
c) Nằm trong khoảng từ $260kg$ đến $270kg$.

13. Chiều cao của 300 sinh viên là một đại lượng ngẫu nhiên có phân phối chuẩn với trung bình $172cm$ và độ lệch tiêu chuẩn $8cm$. Có bao nhiêu sinh viên có chiều cao:

    a) lớn hơn $184cm$,
    b) nhỏ hơn hoặc bằng $160cm$,
    c) giữa $164cm$ và $180cm$,
    d) bằng $172cm$.

14. Cho hai đại lượng ngẫu nhiên độc lập X, Y có bảng phân phối xác suất như sau:

X	1	2	3
P	0,2	0,3	0,5

Y	2	4
P	0,4	0,6

    Tìm phân phối xác suất của $Z = X + Y$.

15. Cho đại lượng ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau:

X	1	3	5
P	0,2	0,5	0,3

    Tìm kỳ vọng và phương sai của đại lượng ngẫu nhiên $Y = \varphi(X) = X^2 + 1$.

16. Gieo một con xúc xắc cân đối n lần. Gọi X là số lần xuất hiện mặt lục. Chứng minh rằng
$$P(\frac{n}{6} - \sqrt{n} < X < \frac{n}{6} + \sqrt{n}) \geq \frac{31}{36}$$

◙ TRẢ LỜI BÀI TẬP

1.

X	0	1	2	3
P	$\frac{5}{30}$	$\frac{15}{30}$	$\frac{9}{30}$	$\frac{1}{30}$

$E(X) = 1,2$, $Var(X) = 0,56$, $mod(X) = 1$.

2.

X	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$E(X) = 7$, $Var(X) = 5,833$.

3. $P(X = 2) = \frac{2}{5}.\frac{1}{4} = \frac{1}{10}$.

$P(X = 3) = \frac{3}{5}.\frac{2}{4}.\frac{1}{3} + \frac{2}{5}.\frac{3}{4}.\frac{1}{3} = \frac{2}{10}$.

$P(X = 4) = \frac{3}{5}.\frac{2}{4}.\frac{2}{3}.\frac{1}{2} + \frac{3}{5}.\frac{2}{4}.\frac{2}{3}.\frac{1}{2} + \frac{2}{5}.\frac{3}{4}.\frac{2}{3}.\frac{1}{2} = \frac{3}{10}$.

$P(X = 5) = 1 - (\frac{2}{20} + \frac{4}{20} + \frac{6}{20}) = \frac{4}{10}$.

Trung bình cần $E(X) = 4$ lần thử.

4. $\dfrac{2}{N}\sum\limits_{i=1}^{n} k_i m_i$.

5. a) Vì $\int\limits_0^4 x^2(4 - x)dx = \frac{64}{3}$ suy ra $k = \frac{3}{64}$, b) $mod(X) = \frac{8}{3}$,

c) $P(X<1)=\frac{3}{64}\int\limits_0^1 x^2(4-x)dx=\frac{13}{256}$.

6. a) $k=4$, b) $F(x)=\begin{cases}1-e^{-2x}(2x^2+2x+1) & \text{nếu } x>0\\ 0 & \text{nếu } x<0\end{cases}$

c) $mod(X)=1$, d) $E(X)=\frac{3}{2}, Var(X)=\frac{3}{4}$.

7. $X\in B(250,2\%)$ a) $P(X=2)=0,0842$, b) $P(x\le 2)=0,1247$.

8. a) $P=0,665$, b) $P=0,619$, c) $P=0,597$.

9. $P(X=x)=\frac{a^x.e^{-a}}{x!}$ với $a=np=(2000).(0,001)=2$.

$P(X=3)=0,18$, $P(X>2)=0,323$.

10. $E(X)=160, Var(X)=19,238$. **11.** $P=0,09$.

12. a) $P(X>300)=1-\phi(1,25)==0,1056$,

b) $P(X,175)==\phi(-1,875)=0,0303$,

c) $P(260<X<270)=\phi(0,5)-\phi(0,25)=0,0928$.

13. a) 18, b) 22, c) 213, d) 14.

14.

Z	3	4	5	6	7
P	0,08	0,12	0,32	0,18	0,3

15. $E(Y)=13,2, Var(Y)=79,36$.

16. X có phân phối nhị thức với $P=\frac{1}{6}$ nên $E(X)=\frac{n}{6}$. Áp dụng bất đẳng thức Tchebyshev ta được bất đẳng thức cần chứng minh.