Xác Suất Thống Kê

Ngày 8 tháng 4 năm 2014

➢ Chương 3. Phân phối xác suất thông dụng

§1. PHÂN PHỐI SIÊU BỘI

1.1. Định nghĩa

- Xét tập có N phần tử gồm N_A phần tử có tính chất A và $N - N_A$ phần tử có tính chất $\overline{A}$. Từ tập đó, ta chọn ra n phần tử.
- Gọi X là số phần tử có tính chất A lẫn trong n phần tử đã chọn thì X có phân phối ***Siêu bội*** (*Hypergeometric distribution*) với 3 tham số N, N_A, n.

Ký hiệu là: $X \in H(N, N_A, n)$ hay $X \sim H(N, N_A, n)$.

➢ Chương 3. Phân phối xác suất thông dụng

- Xác suất trong n phần tử chọn ra có k phần tử A là

$$\boxed{p_k = P(X = k) = \frac{C_{N_A}^{k} C_{N-N_A}^{n-k}}{C_N^n}}$$

Trong đó: $0 \le k \le n$ và $n - (N - N_A) \le k \le N_A$.

VD 1. Một hộp phấn gồm 10 viên, trong đó có 7 viên màu trắng. Lấy ngẫu nhiên 5 viên phấn từ hộp này. Gọi X là số viên phấn trắng lấy được.
Lập bảng phân phối xác suất và tính kỳ vọng của X?

Giải. Ta có: $X = \{2;\ 3;\ 4;\ 5\}$ và

$$N = 10,\ N_A = 7,\ n = 5 \Rightarrow X \in H(10,\ 7,\ 5).$$

➤ Chương 3. Phân phối xác suất thông dụng

- Bảng phân phối xác suất của X là

X	2	3	4	5
P	$\frac{C_7^2 C_3^3}{C_{10}^5}$	$\frac{C_7^3 C_3^2}{C_{10}^5}$	$\frac{C_7^4 C_3^1}{C_{10}^5}$	$\frac{C_7^5 C_3^0}{C_{10}^5}$

- $EX = 2\frac{C_7^2 C_3^3}{C_{10}^5} + 3\frac{C_7^3 C_3^2}{C_{10}^5} + 4\frac{C_7^4 C_3^1}{C_{10}^5} + 5\frac{C_7^5 C_3^0}{C_{10}^5} = \frac{7}{2}.$

1.2. Các số đặc trưng của $X \sim H(N, N_A, n)$

$$\boxed{EX = np; \quad VarX = npq\frac{N-n}{N-1}}$$

trong đó: $p = \frac{N_A}{N}, \quad q = 1 - p.$

➢ Chương 3. Phân phối xác suất thông dụng

VD 2. Một cửa hàng bán 100 bóng đèn, trong đó có 12 bóng hỏng. Một người chọn mua ngẫu nhiên 15 bóng đèn từ cửa hàng này. Hỏi trung bình người đó mua được bao nhiêu bóng đèn tốt ?

VD 3. Tại một công trình có 100 người đang làm việc, trong đó có 70 kỹ sư. Chọn ngẫu nhiên 40 người từ công trình này. Gọi X là số kỹ sư chọn được.

1) Tính xác suất chọn được từ 27 đến 29 kỹ sư ?
2) Tính EX và $VarX$?

……………………………………………………………

➢ Chương 3. Phân phối xác suất thông dụng

§2. PHÂN PHỐI NHỊ THỨC

2.1. Phân phối Bernoulli

a) Định nghĩa

- Phép thử Bernoulli là một phép thử mà ta chỉ quan tâm đến 2 biến cố A và $\overline{A}$, với $P(A) = p$.
- Xét biến ngẫu nhiên:

$$X = \begin{cases} 1 \; khi \; A \; xảy \; ra, \\ 0 \; khi \; \overline{A} \; xảy \; ra, \end{cases} \quad P(\overline{A}) = 1 - p = q.$$

Khi đó, ta nói X có phân phối Bernoulli với tham số p.

Ký hiệu là $X \in B(p)$ hay $X \sim B(p)$.

Bảng phân phối xác suất của X là:

X	0	1
P	q	p

➢ Chương 3. Phân phối xác suất thông dụng

b) Các số đặc trưng của $X \sim B(p)$

$$\boxed{EX = p;\ \ VarX = pq}$$

VD 1. Một câu hỏi trắc nghiệm có 4 phương án trả lời, trong đó chỉ có 1 phương án đúng. Một sinh viên chọn ngẫu nhiên 1 phương án để trả lời câu hỏi đó.

Gọi A: "sinh viên này trả lời đúng".

Khi đó, việc trả lời câu hỏi của sinh viên này là một phép thử Bernoulli và $p = P(A) = 0,25$, $q = 0,75$.

Gọi BNN $X = \begin{cases} 1 \ khi\ sinh\ viên\ này\ trả\ lời\ đúng, \\ 0 \ khi\ sinh\ viên\ này\ trả\ lời\ sai, \end{cases}$

thì $X \in B(0,25)$ và $EX = 0,25$, $VarX = 0,1875$.

➤ Chương 3. Phân phối xác suất thông dụng

2.2. Phân phối Nhị thức

a) Định nghĩa

- Xét dãy n phép thử Bernoulli độc lập. Với phép thử thứ i, ta xét biến ngẫu nhiên $X_i \in B(p)$ $(i = 1, \ldots, n)$.

Nghĩa là, $X_i = \begin{cases} 1 \; khi \; lần \; thứ \; i \; A \; xảy \; ra, \\ 0 \; khi \; lần \; thứ \; i \; A \; không \; xảy \; ra. \end{cases}$

- Gọi X là số lần biến cố A xuất hiện trong n phép thử. Khi đó, $X = X_1 + \ldots + X_n$ và ta nói X có phân phối *Nhị thức* (*Binomial distribution*) với tham số n, p. Ký hiệu là $X \in B(n,\, p)$ hay $X \sim B(n,\, p)$.

➢ Chương 3. Phân phối xác suất thông dụng

- Xác suất trong n lần thử có k lần A xảy ra là

$$\boxed{p_k = P(X = k) = C_n^k p^k q^{n-k} \quad (k = 0, 1, \ldots, n)}$$

VD 2. Một đề thi XSTK gồm 20 câu hỏi trắc nghiệm như trong VD 1. Sinh viên B làm bài một cách ngẫu nhiên. Biết rằng, nếu trả lời đúng 1 câu thì sinh viên B được 0,5 điểm và nếu trả lời sai 1 câu thì bị trừ 0,125 điểm. Tính xác suất để sinh viên B đạt điểm 5 ?

b) Các số đặc trưng của $X \sim B(n, p)$

$$EX = np; \quad VarX = npq$$

$$\text{Mod}\, X = x_0 \in \mathbb{N}:\ np - q \le x_0 \le np - q + 1$$

➢ Chương 3. Phân phối xác suất thông dụng

VD 3. Ông B trồng 100 cây bạch đàn với xác suất cây chết là 0,02. Gọi X là số cây bạch đàn chết.

1) Tính xác suất có từ 3 đến 5 cây bạch đàn chết ?

2) Tính trung bình số cây bạch đàn chết và $VarX$?

3) Hỏi ông B cần phải trồng tối thiểu mấy cây bạch đàn để xác suất có ít nhất 1 cây chết lớn hơn 50% ?

VD 4. Một nhà vườn trồng 126 cây lan quý, xác suất nở hoa của mỗi cây trong 1 năm là 0,67.

1) Giá bán 1 cây lan quý nở hoa là 2 triệu đồng. Giả sử nhà vườn bán hết những cây lan nở hoa thì mỗi năm nhà vườn thu được chắc chắn nhất là bao nhiêu tiền?

2) Nếu muốn trung bình mỗi năm có nhiều hơn 100 cây lan quý nở hoa thì nhà vườn phải trồng tối thiểu mấy cây lan quý ?

➢ Chương 3. Phân phối xác suất thông dụng

VD 5. Có 10 hộp phấn màu giống nhau, mỗi hộp chứa 20 viên phấn gồm hai loại: 3 hộp loại I, mỗi hộp có 12 viên phấn đỏ; 7 hộp loại II, mỗi hộp có 8 viên phấn đỏ. Chọn ngẫu nhiên 1 hộp và từ hộp đó lấy lần lượt ra 5 viên phấn (lấy viên nào xong thì trả lại vào hộp viên đó). Tính xác suất chọn được 3 viên phấn đỏ ?

VD 6. Một lô hàng chứa 20 sản phẩm trong đó có 4 phế phẩm. Chọn liên tiếp 3 lần từ lô hàng (mỗi lần chọn có hoàn lại), mỗi lần chọn ra 4 sản phẩm. Tính xác suất để trong 3 lần chọn có ít nhất 1 lần chọn phải 2 phế phẩm ?

……………………………………………………………………

➢ Chương 3. Phân phối xác suất thông dụng

§3. PHÂN PHỐI POISSON

3.1. Bài toán dẫn đến phân phối Poisson

- Giả sử các vụ tai nạn giao thông ở vùng A xảy ra một cách ngẫu nhiên, độc lập với nhau và trung bình 1 ngày có λ vụ tai nạn. Gọi X là số vụ tai nạn giao thông xảy ra trong 1 ngày ở vùng A.
- Chia 24 giờ trong ngày thành n khoảng thời gian sao cho ta có thể coi rằng trong mỗi khoảng thời gian đó có nhiều nhất 1 vụ tai nạn xảy ra, và khả năng xảy ra tai nạn giao thông trong mỗi khoảng thời gian bằng $\dfrac{\lambda}{n}$.

Khi đó, $X \in B\left(n, \dfrac{\lambda}{n}\right)$.

➢ Chương 3. Phân phối xác suất thông dụng

- Ta có: $P(X=k)=C_n^k\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$

$$=\frac{n!}{k!(n-k)!}\cdot\frac{\lambda^k}{n^k}\cdot\frac{1}{(n-\lambda)^k.n^{-k}}\cdot\left(1-\frac{\lambda}{n}\right)^n$$

$$=\frac{\lambda^k}{k!}\cdot\frac{n(n-1)\dots(n-k+1)}{(n-\lambda)^k}\cdot\left(1-\frac{\lambda}{n}\right)^n.$$

Suy ra:

$$P(X=k)\xrightarrow{n\to\infty}\frac{\lambda^k}{k!}.e^{-\lambda}.$$

> Chương 3. Phân phối xác suất thông dụng

3.2. Định nghĩa phân phối Poisson

Biến ngẫu nhiên X được gọi là có phân phối Poisson tham số $\lambda > 0$, ký hiệu là $X \in P(\lambda)$ hay $X \sim P(\lambda)$, nếu X nhận các giá trị $0, 1, 2, \ldots, n, \ldots$ với xác suất

$$\boxed{p_k = P(X = k) = \frac{e^{-\lambda}.\lambda^k}{k!} \quad (k = 0, 1, \ldots, n, \ldots)}$$

λ là trung bình số lần xuất hiện biến cố ta quan tâm trong một khoảng xác định (khoảng thời gian hoặc khoảng đơn vị tính).

VD. Quan sát tại một sân bay thấy trung bình 16 phút có 2 máy bay hạ cánh. Suy ra trong 1giờ trung bình có:

➢ Chương 3. Phân phối xác suất thông dụng

$$\lambda = \frac{60.2}{16} = 7,5 \text{ máy bay hạ cánh.}$$

VD. Trung bình cứ 100 sinh viên thi hết môn XSTK có 71 sinh viên thi đạt. Suy ra 120 sinh viên thi hết môn XSTK thì trung bình có 85,2 sinh viên thi đạt.

3.3. Các số đặc trưng của $X \sim P(\lambda)$

$$\boxed{\begin{array}{l} EX = VarX = \lambda \\ \text{Mod}\, X = x_0 \in \mathbb{N}: \lambda - 1 \le x_0 \le \lambda \end{array}}$$

VD 1. Quan sát tại siêu thị A thấy trung bình 5 phút có 18 khách đến mua hàng.

1) Tính xác suất để trong 7 phút có 25 khách đến siêu thị A ?

➢ Chương 3. Phân phối xác suất thông dụng

2) Tính xác suất để trong 2 phút có từ 3 đến 5 khách đến siêu thị A ?

3) Tính số khách chắc chắn nhất sẽ đến siêu thị A trong 1 giờ ?

VD 2. Quan sát thấy trung bình 2 phút có 6 ôtô đi qua trạm thu phí. Biết xác suất có ít nhất 1 ôtô đi qua trạm thu phí trong t phút bằng 0,9. Giá trị của t (phút) là:

A. 0,9082 B. 0,8591 C. 0,8514 D. 0,7675.

VD 3. Cứ mỗi lần đi câu cá thì ông A chọn ngẫu nhiên 1 trong 2 nơi để câu. Nếu đi câu ở địa điểm I thì trung bình cứ 10 lần móc mồi, ông A câu được 2 con cá; câu ở địa điểm II thì trung bình cứ 12 lần móc mồi, ông A câu được 3 con cá.

➢ Chương 3. Phân phối xác suất thông dụng

Hôm nay ông A đi câu, ông đã móc mồi 20 lần và câu được 5 con cá. Tính xác suất ông A câu được 5 con cá ở địa điểm II ?

VD 4. Tại một xưởng dệt, trung bình dệt 10 m vải loại B thì bị lỗi 13 chỗ. Chọn lần lượt 5 xấp vải loại B của xưởng, mỗi xấp dài 6 m. Tính xác suất để 3 trong 5 xấp vải ấy, mỗi xấp vải có đúng 7 chỗ bị lỗi ?

VD 5. Quan sát thấy trung bình 1 ngày (24 giờ) có 12 chuyến tàu vào cảng A. Chọn ngẫu nhiên 6 giờ trong 1 ngày. Tính xác suất để 2 trong 6 giờ ấy, mỗi giờ có đúng 1 tàu vào cảng A ?

……………………………………………………………………

➤ Chương 3. Phân phối xác suất thông dụng

§4. PHÂN PHỐI CHUẨN

4.1. Phân phối chuẩn

a) Định nghĩa

Biến ngẫu nhiên liên tục X được gọi là có phân phối chuẩn (*Normal distribution*) với hai tham số μ và σ^2 $(\sigma > 0)$, ký hiệu là $X \in N(\mu;\ \sigma^2)$ hay $X \sim N(\mu;\ \sigma^2)$, nếu hàm mật độ xác suất của X có dạng

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ x \in \mathbb{R}}$$

➢ Chương 3. Phân phối xác suất thông dụng

b) Các số đặc trưng của $X \sim N(\mu, \sigma^2)$

$$\boxed{\text{Mod}\,X = EX = \mu;\;\; VarX = \sigma^2}$$

c) Xác suất của $X \sim N(\mu, \sigma^2)$

$$\boxed{P(a \le X \le b) = \int_a^b f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}$$

Nhận xét. Đổi biến $z = \dfrac{x-\mu}{\sigma}$, ta có:

$$\frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.$$

➢ Chương 3. Phân phối xác suất thông dụng

4.2. Phân phối chuẩn tắc

a) Định nghĩa

Biến ngẫu nhiên Z có phân phối chuẩn với hai tham số $\mu = 0$ và $\sigma^2 = 1$ được gọi là có phân phối chuẩn tắc (hay *phân phối* ***Gauss***), ký hiệu là $Z \in N(0;\ 1)$ hay $Z \sim N(0;\ 1)$. Hàm mật độ xác suất của Z là

$$\boxed{f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}},\ z \in \mathbb{R}}$$

(Giá trị hàm $f(z)$ được cho trong bảng phụ lục A).