Phân Phối Nhị Thức

§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:

$$X = \begin{cases} 1 \ \textit{khi } A \textit{ xuất hiện}, \\ 0 \ \textit{khi } \overline{A} \textit{ xuất hiện}, \end{cases} \quad P(\overline{A}) = 1 - p = q.$$

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

b) Các số đặc trưng của $X \sim B(p)$

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

<u>VD 1.</u> 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 $p = P(A) = \dfrac{1}{4}$, $q = \dfrac{3}{4}$.

Gọi BNN $X = \begin{cases} 1 \text{ khi sinh vieân naøy traû lôøi ñuùng,} \\ 0 \text{ khi sinh vieân naøy traû lôøi sai,} \end{cases}$

thì $X \in B\left(\dfrac{1}{4}\right)$ và $EX = \dfrac{1}{4}$, $VarX = \dfrac{1}{4}.\dfrac{3}{4} = \dfrac{3}{16}$.

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: $$X_i = \begin{cases} 1 \text{ khi laàn thöû } A \text{ xuaát hieän,} \\ 0 \text{ khi laàn thöû } \bar{A} \text{ xuaát hieän.} \end{cases}$$

- 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 $X \sim B(n,\ p)$.

- Xác suất trong n lần thử có k lần A xuất hiện 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 ?

Giải. Gọi x là số câu đúng để sinh viên B đạt điểm 5.

Ta có: $0,5.x - 0,125(20 - x) = 5 \Rightarrow x = 12$ (câu).

Vậy xác suất để sinh viên B đạt điểm 5 là:

$$p_5 = C_{20}^{12}(0,25)^{12}(0,75)^8 \approx 0,0008.$$

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

$$\boxed{\begin{array}{c} EX = np; \;\; VarX = npq; \\ ModX = x_0 : np - q \le x_0 \le np - q + 1. \end{array}}$$

<u>VD 3.</u> 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 10% ?

Giải

1) Ta có: $X = \{0,\ 1,\ 2, \ldots, 100\}$ và

$$n = 100,\ p = 0,02 \Rightarrow X \in B(100;\ 0,02).$$

$$P(3 \le X \le 5) = p_3 + p_4 + p_5$$

$$= C_{100}^{3}(0,02)^3(0,98)^{97} + C_{100}^{4}(0,02)^4(0,98)^{96}$$

$$+C_{100}^{5}(0,02)^5(0,98)^{95} = 0,3078.$$

2) Trung bình số cây bạch đàn chết là:

$$EX = np = 100.0,02 = 2 \text{ (cây)}.$$

Phương sai là: $VarX = npq = 1,96.$

3) Gọi N là số cây bạch đàn ông B cần phải trồng.

Ta có:

$$p(X \geq 1) = 1 - P(X = 0) = 1 - p_0$$

$$= 1 - C_N^0 (0,02)^0 (0,98)^N = 1 - (0,98)^N.$$

$$P(X \geq 1) > 0,1 \Leftrightarrow 1 - (0,98)^N > 0,1$$

$$\Leftrightarrow (0,98)^N < 0,9$$

$$\Leftrightarrow N \ln 0,98 < \ln 0,9$$

$$\Leftrightarrow N > 5,2152.$$

Vậy ông B cần phải trồng ít nhất 6 cây bạch đàn.

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 ?

Giải. Gọi X là số cây lan quý nở hoa trong 1 năm.

$X = \{0, 1, 2, ..., 126\}, \ p = 0,67 \Rightarrow X \in B(126;\ 0,67).$

1) Ta có $ModX = x_0$:

$$np - q \le x_0 \le np - q + 1 \Rightarrow 84,09 \le x_0 \le 85,09$$

$\Rightarrow x_0 = 85$ là số cây lan chắc chắn nhất sẽ nở hoa.

Vậy nhà vườn thu được $85.2 = 170$ triệu đồng.

2) Gọi N là số cây lan quý nhà vườn phải trồng, ta có:

$$EX > 100 \Rightarrow N.0,67 > 100 \Rightarrow N > 149,25.$$

Vậy nhà vườn phải trồng tối thiểu 150 cây.

VD 5. Một nhà tuyển dụng kiểm tra kiến thức lần lượt các ứng viên, xác suất được chọn của mỗi ứng viên đều bằng 0,56. Biết xác suất để nhà tuyển dụng chọn đúng 8 ứng viên là 0,0843. Số người cần phải kiểm tra là:

A. 9 người; B. 10 người;

C. 12 người; D. 13 người.

Giải. Gọi n, $n \geq 8$ là số người cần phải kiểm tra và X là số người được chọn, ta có:

$$X = \{0;\, 1;\, 2; \ldots;\, n\},\ p = 0,56 \Rightarrow X \in B(n;\, 0,56)$$

$$\Rightarrow P(X = 8) = C_n^8 (0,56)^8 (0,44)^{n-8}.$$

Thử trực tiếp từ đáp án, ta được đáp án đúng là B.

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

<u>Giải</u>

- Gọi X là số phế phẩm chọn phải, ta có:

$$X = \{0; 1; \dots; 4\}, \; X \in H(20;\, 4;\, 4).$$

Xác suất chọn phải 2 phế phẩm là:

$$P(X = 2) = \frac{C_4^2 C_{16}^2}{C_{20}^4} = 0,1486.$$

- Gọi Y là số lần chọn phải 2 phế phẩm trong 3 lần.

Do chọn có hoàn lại nên mỗi lần chọn là độc lập với xác suất $p = 0,1486 \Rightarrow Y \in B(3;\ 0,1486)$.

Vậy xác suất trong 3 lần chọn có 1 lần đạt yêu cầu là:

$$p_1 = C_3^1.0,1486.(1-0,1486)^2 = 0,3232.$$

....