Sơ lược

Lý thuyết xác suất và thống kê toán 1

I. Chương I: Biến cố và xác suất biến cố

1. Biến cố

1.1.Định nghĩa

Hiện tượng có thể xảy ra (hoặc không xảy ra) trong kết quả của phép thử gọi là biến cố.

1.2.Phân loại và kí hiệu biến cố

- *Biến cố chắc chắn*: Là hiện tượng chắc chắn phải xảy ra trong 1 phép thử. Kí hiệu U
- *Biến cố không thể có*: Là hiện tượng không thể xảy ra trong 1 phép thử. Kí hiệu là V
- *Biến cố ngẫu nhiên*: Là hiện tượng mà có thể xảy ra hoặc không sau phép thử. Kí hiệu là A, B, C, D, …

2. Xác suất biến cố

2.1.Định nghĩa

- Để đặc trưng cho sự xảy ra hay không xảy ra (một cách khách quan) của biến cố, người ta dùng một giá trị số thực gọi là xác suất của biến cố.

2.2.Các tính chất: xác suất của biến cố A bất kì có cách tính chất

- $0 \le P(A) \le 1$, $P(U)=1$ $P(V)=0$

U: biến cố chắc chắn
V: biến cố ko thể xảy ra

2.3.Nguyên lý xác suất lớn, nguyên lý xác suất nhỏ

- *Nguyên lý xác suất lớn:*

  + Nếu một biến cố có xác suất rất nhỏ thực tế có thể cho rằng trong một phép thử biến cố đó sẽ không xảy ra.
  + Mức xác suất được coi là nhỏ tùy thuộc vào từng bài toán và gọi là mức ý nghĩa.
  + Nguyên lý XS nhỏ là cơ sở của phương pháp kiểm định.
- *Nguyên lý xác suất lớn:*

  +Nếu một biến cố có xác suất rất lớn thực tế có thể cho rằng trong một phép thử biến cố đó sẽ xảy ra.
  + Mức xác suất đủ lớn gọi là độ tin cậy.
  + Nguyên lý XS lớn là cơ sở của phương pháp ước lượng bằng khoảng tin cậy.

3. Các phương pháp xác định xác suất biến cố

3.1.Dùng định nghĩa cổ điển

- Xác suất của biến cố A được xác định bởi: $P(A)=\dfrac{M}{N}$, trong đó: N là tổng số kết cục duy nhất đồng khả năng(không gian mẫu) khi thực hiện phép thử, M là số kết cục thuận lợi cho A (số kết cục mà biến cố A có thể xảy ra)

3.2.Dùng công thức giải tích tổ hợp

- Mục đích là để hỗ trợ việc đếm các kết cục cho thuận tiện hơn, ta có thể sử dụng:
- Hoán vị của n phần tử: n! là số cách xắp xếp ngẫu nhiên n phần tử theo 1 thứ tự nào đó.
- Chỉnh hợp chập k của n phần tử: A_n^k là số cách lấy ngẫu nhiên (***không hoàn lại***) ra k phần tử bất kì trong n phần tử ***theo 1 thứ tự định trước***.
- Chỉnh hợp lặp chập k của n phần tử: $\overline{A}_n^k$ là số cách lấy ngẫu nhiên (***có hoàn lại***) ra k phần tử bất kì trong n phần tử.
- Tổ hợp chập k của n phần tử: C_n^k là số cách lấy ngẫu nhiên (***không hoàn lại***) ra k phần tử bất kì trong n phần tử.

3.3.*Dùng sơ đồ Ven* (Cụ thể xem giáo trình)

- Tùy vào dữ kiện, vẽ phác họa phạm vi các biến cố bằng vòng tròn. Khu vực 2 vòng tròn giao nhau là khu vực 2 biến cố cùng đồng thời xảy ra.
- Điền giá trị xác suất, ưu tiên điền phần giao nhau trước, rồi đến cách phần còn lại.

3.4.Định nghĩa thống kê về xác suất

- Thực hiện n phép thử thấy biến cố A xuất hiện m lần, tần suất của biến cố A là $f(A) = \dfrac{m}{n}$
- Bằng thực nghiệm, người ta chứng minh được rằng: $\lim\limits_{n \to +\infty} f(A) = P(A)$

4. Định lý cộng

4.1.Biến cố tổng

- Biến cố C = A + B xảy ra khi và chỉ khi có ít nhất 1 trong 2 biến cố A, B xảy ra.

4.2.Biến cố xung khắc và xác suất tổng 2 biến cố xung khắc

- A, B là 2 biến cố xung khắc thì A và B không thể cùng xảy ra trong 1 phép thử. Lúc đó:
$$P(A+B) = P(A) + P(B)$$
- Nhóm n biến cố $A_i\left(i = \overline{1,n}\right)$ gọi là xung khắc từng đôi nếu 2 biến cố bất kì trong chúng không thể đồng thời xảy ra trong 1 phép thử. Lúc đó: $P\left(\sum\limits_{i=1}^{n} A_i\right) = \sum\limits_{i=1}^{n} P(A_i)$

5. Định lý nhân

5.1.Biến cố tích

- Biến cố C = A.B xảy ra khi và chỉ khi đồng thời 2 biến cố A và B xảy ra

5.2.Biến cố độc lập, biến cố phụ thuộc

- Biến cố A độc lập với biến cố B khi và chỉ khi xác suất xảy ra A không ảnh hưởng tới xác suất xảy ra B và ngược lại. Lúc ấy ta có: $P(AB) = P(A)P(B)$
- Nếu xác suất của chúng ảnh hưởng nhau, thì A và B là hai biến cố phụ thuộc. Lúc này xác suất của A tính trong điều kiện B đã xảy ra được xác định bởi: $P(A/B) = \dfrac{P(AB)}{P(B)};\left(P(B) \neq 0\right)$

5.3.*Xác suất của tổng 2 biến cố không xung khắc:* $P(A+B) = P(A) + P(B) - P(AB)$

5.4.Nhóm đầy đủ các biến cố - Biến cố đối

- Nhóm các biến cố $H_i\left(i = \overline{1,n}\right)$ tạo thành ***nhóm đầy đủ*** các biến cố nếu mỗi biến cố đều là các khả năng duy nhất của phép thử, đồng thời thỏa mãn:

$$\sum_{i=1}^{n} P(H_i) = P(H_1) + P(H_2) + ... + P(H_n) = 1$$

- ***Biến cố đối*** của A được kí hiệu là $\overline{A}$, là hai biến cố tạo nên 1 nhóm đầy đủ các biến cố. Lưu ý rằng, biến cố đối thì sẽ xung khắc, nhưng biến cố xung khắc thì chưa chắc là đối.

6. Hệ quả của định lý cộng và định lý nhân

6.1.Công thức Bernoulli

- Lược đồ Bernoulli:

  (+) Có n phép thử độc lập

  (+) Trong mỗi phép thử, biến cố A nào đó xảy ra hoặc không với $P(A) = p$

- Lúc đó xác suất để A xảy ra k lần được xác định bởi công thức: $P_k = C_n^k p^k (1-p)^{n-k}$

6.2.Công thức xác suất đầy đủ, bayes

- Biến cố A xảy ra với 1 trong các biến cố của nhóm đầy đủ $H_i\left(i=\overline{1,n}\right)$, lúc đó:

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A/H_i) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) + ... + P(H_n)P(A/H_n)$$

- Công thức Bayes nhằm đánh giá lại xác suất xảy ra của các biến cố trong nhóm đầy đủ khi biết rằng A đã xảy ra: $P(H_i / A) = \dfrac{P(H_i)P(A/H_i)}{P(A)}$

  Về bản chất nó chính là khai triển xác suất có điều kiện nên không cần nhớ cũng được!

7. Một số biến đổi đặc biệt cần nhớ

- Với A bất kì, U là biến cố không thể có:

$$A.U = A;\ AV = V;\ A+V = A;\ A+U = U;\ A.A = A;\ A.\overline{A} = V;\ A+\overline{A} = U$$

- Với A, B, C bất kì, ta có:

  (1) $P(A+B+C) = P(A) + P(B+C) - P(AB+AC) = P(A) + \left[P(B) + P(C) - P(BC)\right] - \left[P(AB) + P(BC) - P(ACBC)\right]$

  $= P(A) + P(B) + P(C) - P(AB) - P(BC) - P(CA) + P(ABC)$

  (2) $P(ABC) = P\left[A(BC)\right] = P(A)P(BC/A) = P(A)P(B/A)P(C/AB)$

  (3) $P(A/B) + P(\overline{A}/B) = 1,\ P(AB) + P(A\overline{B}) = P(A)P(B/A) + P(A)P(\overline{B}/A) = P(A)$

  (4) $P(A/\overline{B}) = \dfrac{P(A\overline{B})}{P(\overline{B})} = \dfrac{P(A)P(\overline{B}/A)}{1-P(B)} = \dfrac{P(A)-P(AB)}{1-P(B)}$

8. Gợi ý các dạng toán thường gặp

8.1.Dạng 1: *Sử dụng thuần túy định nghĩa cổ điển (Khó nhất)*

8.2.Dạng 2: *Sử dụng trực tiếp định lý cộng và định lý nhân*

8.3.Dạng 3: *Sử dụng công thức Bernoulli*

8.4.Dạng 4: *Công thức xác suất đầy đủ và các kết hợp*

II. Chương II: Biến ngẫu nhiên và quy luật phân phối xác suất

1. Khái niệm biến ngẫu nhiên

1.1. Định nghĩa

❖ Biến ngẫu nhiên là một biến số thực, nhận các giá trị có thể có của nó một cách ngẫu nhiên sau khi thực hiện phép thử tương ứng.

1.2.Phân loại biến ngẫu nhiên

❖ Ngẫu biến rời rạc: nhận các giá trị trong 1 tập hợp rời rạc.

❖ Ngẫu biến liên tục: nhận các giá trị trong 1 tập hợp liên tục.

❖ Các ngẫu biến thường kí hiệu là X, Y, Z, T, W, …

2. Quy luật phân phối xác suất của biến ngẫu nhiên

Là cách thức biểu diễn giá trị của ngẫu biến tương ứng với xác suất để ngẫu biến ngận giá trị đó.

2.1.Bảng phân phối xác suất

❖ Bảng phân phối xác suất của X có dạng:

X	x_1	x_2	x_3	$\cdots$	x_n
Xác suất	p_1	p_2	p_3	$\cdots$	p_n

Trong đó:

(1) $x_i\left(i=\overline{1,n}\right)$ là các giá trị có thể có của biến ngẫu nhiên, được xếp tăng dần từ bên trái.

(2) $p_i = P(X = x_i),\ \left(i=\overline{1,n}\right)$ và $\sum_{i=1}^{n} p_i = p_1 + p_2 + ... + p_n = 1$

❖ Một số công thức xác suất cần lưu ý:

$P(x_1 \le X \le x_3) = P\left(X = x_1\right) + P\left(X = x_2\right) + P(X = x_3)$

$P(x_1 < X \le x_3) = P\left(X = x_2\right) + P(X = x_3)$; $P(x_1 \le X < x_3) = P(X = x_1) + P\left(X = x_2\right)$

$P(X \le x_n) = 1 = P\left(X < +\infty\right)$; $P(X < x_1) = 0 = P\left(X < -\infty\right)$

2.2.Hàm phân bố xác suất

❖ Định nghĩa: $F(x) = P(X < x)$ là giá trị cho biết mức độ xác suất ở bên trái số thực *x*.

❖ Tính chất:

(1) $0 \le F(x) \le 1$ và *F(x)* là hàm không giảm theo *x*

(2) $F(+\infty) = P(X < +\infty) = 1,\ F(-\infty) = P\left(X < -\infty\right) = 0$

(3) $P(a \le X < b) = P(X < b) - P\left(X < a\right) = F(b) - F(a)$

2.3.Hàm mật độ xác suất

❖ Định nghĩa: $f(x) = F'(x)$ => chỉ dùng cho ngẫu biến liên tục.

❖ Tính chất:

(1) $f(x) \ge 0$ (Do *F(x)* là hàm không giảm)

(2) $P\left(a \le X < b\right) = \int_a^b f(x)dx = F\left(b\right) - F\left(a\right)$; $\int_{-\infty}^{+\infty} f(x)dx = 1$

(3) Với ngẫu biến liên tục X thì: $P(X = a) = 0$ nên:

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$$

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

3. Các tham số đặc trưng của biến ngẫu nhiên

3.1.*Kì vọng* $E(X)$

- Là giá trị trung bình của X, phản ánh xu hướng trung tâm của phân phối xác suất.
- Xác định bởi:

  (1) Ngẫu biến rời rạc: $E(X)=\sum p_i x_i = p_1x_1 + p_2x_2 + p_3x_3 + ... + p_nx_n$

  (2) Ngẫu biến liên tục: $E(X)=\int_{-\infty}^{+\infty} xf(x)dx$

- Tính chất:
  - ✓ $E(C)=C$; $E(CX)=CE(X)$ với C là hằng số
  - ✓ $E(aX+bY)=aE(X)+bE(Y)$: a, b là hằng số; X và Y là 2 ngẫu biến bất kì.
  - ✓ $E(XY)=E(X)E(Y)$ nếu X và Y là hai biến ngẫu nhiên độc lập. Độc lập tức là quy luật phân phối xác suất của X không ảnh hưởng tới quy luật phân phối của Y.

3.2.*Phương sai* $V(X)$

- Là trung bình tổng của bình phương các sai lệch giữa X và $E(X)$. Nó đặc trưng cho sự phân tán (biến động, đồng đều hay ổn định) của phân phối xác suất.
- Xác định bởi: $V(X)=E\left[\left(X-E(X)\right)^2\right]=E\left(X^2\right)-E^2(X)$

  (1) Ngẫu biến rời rạc: $V(X)=p_1x_1^2+p_2x_2^2+p_3x_3^2+...+p_nx_n^2-E^2(X)$

  (2) Ngẫu biến liên tục: $V(X)=\int_{-\infty}^{+\infty} x^2 f(x)dx - E^2(X)$

- Tính chất:
  - ✓ $V(C)=0$; $V(CX)=C^2V(X)$ với C là hằng số.
  - ✓ $V(aX+bY)=a^2V(X)+b^2V(Y)$ nếu X độc lập Y
- Phương sai có đơn vị đo bằng ***BÌNH PHƯƠNG*** đơn vị của biến ngẫu nhiên (X).

3.3.*Độ lệch chuẩn* $\sigma=\sqrt{V(X)}$ về ý nghĩa giống hệt phương sai, chỉ khác đơn vị đo.

3.4.*Mốt* m_0: là giá trị có xác suất lớn nhất (có khả năng xảy ra cao nhất) trong phân phối.

3.5.*Hệ số biến thiên, hệ số nhọn, hệ số bất đối xứng* (Ít gặp, xem giáo trình)

4. Gợi ý các dạng bài tập thường gặp

4.1.*Dạng 1*: Bài toán lập bảng

4.2.*Dạng 2*: Bài toán xử lý thông tin từ bảng

(1) *Tính kì vọng, phương sai, mốt và nêu ý nghĩa*

(2) *Tính các giá trị xác suất*

(3) *Dạng toán ra quyết định trong kinh tế bằng kì vọng và phương sai (rủi ro)*

(4) *Các dạng khác*

III. Chương III: Một số quy luật phân phối xác suất thông dụng

1. Phân phối Không-Một và phân phối Nhị Thức

1.1.Phân phối Không-Một $X \sim A(p)$

- Giả thiết: lược đồ Bernoulli. Khi lược đồ được thỏa mãn, nếu gọi X là số lần biến cố A đó xảy ra
- Các tham số đặc trưng:
  (1) Kì vọng: $E(X) = p$; (2) Phương sai: $V(X) = p(1-p)$

1.2.Phân phối Nhị Thức $X \sim B(n;p)$

- Giả thiết: lược đồ Bernoulli
- Các tham số đặc trưng:
  (1) *Kì vọng*: $E(X) = np$
  (2) *Phương sai*: $V(X) = np(1-p)$
  (3) *Mốt*: $np + p - 1 \le m_0 \le np + p$
- Chú ý: Về bản chất, biến ngẫu nhiên phân phối Nhị Thức là tổng của n biến ngẫu nhiên phân phối Không-Một.
- Tần suất $f = \dfrac{X}{n}$ cũng phân phối $B(n;p)$, với: $E(f) = p$, $V(f) = \dfrac{p(1-p)}{n}$

2. Phân phối lũy thừa $X \sim E(\lambda)$

2.1.Hàm mật độ $f(x) = \begin{cases} 0 & khi\ x < 0 \\ \lambda e^{-\lambda x} & khi\ x \ge 0 \end{cases}$

2.2.Các tham số đặc trưng: $E(X) = \dfrac{1}{\lambda}$; $V(X) = \dfrac{1}{\lambda^2}$

3. Phân phối Chuẩn $X \sim N(\mu;\sigma^2)$ **và phân phối chuẩn hóa** $U \sim N(0;1)$

3.1.Hàm mật độ của phân phối chuẩn (SGT)

3.2.Các tham số đặc trưng $E(X) = \mu$; $V(X) = \sigma^2$

3.3.Hàm phân bố chuẩn hóa

- Nếu biến ngẫu nhiên $X \sim N(\mu;\sigma^2)$ thì ta có thể chuẩn hóa nó bằng cách đặt:

$$U = \frac{X-\mu}{\sigma}$$

  Lúc đó:

$$E(U) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{E(X)-\mu}{\sigma} = 0;\ V(U) = V\left(\frac{X-\mu}{\sigma}\right) = \frac{V(X-\mu)}{\sigma^2} = \frac{V(X)}{\sigma^2} = 1$$

  U sẽ phân phối chuẩn hóa. Kí hiệu: $U \sim N(0;1)$

- Hàm phân bố chuẩn hóa $\Phi(u)$ và hàm $\Phi_0(u)$ có liên hệ:

$$\Phi(u) = \int_{-\infty}^{u} \phi(u)du = \int_{-\infty}^{0} \phi(u)du + \int_{0}^{u} \phi(u)du = 0,5 + \int_{0}^{u} \phi(u)du = 0,5 + \Phi_0(u)$$

  Trong đó $\phi(u)$ là hàm mật độ chuẩn hóa, và $\Phi_0(u)$ có các tính chất quan trọng:

  (1) $\Phi_0(-u) = -\Phi_0(u)$

(2) $\Phi_0(u) = 0,5$ với mọi $u \geq 5$

3.4.Các công thức phải nhớ:

(1) $P(a \leq X \leq b) = \Phi_0\left(\dfrac{b-\mu}{\sigma}\right) - \Phi_0\left(\dfrac{a-\mu}{\sigma}\right)$

(2) $P(X \geq a) = 0,5 - \Phi_0\left(\dfrac{a-\mu}{\sigma}\right)$ và $P(X \leq b) = 0,5 + \Phi_0\left(\dfrac{b-\mu}{\sigma}\right)$

(3) $P\left(\left|X - E(X)\right| \leq \varepsilon\right) = 2\Phi_0\left(\dfrac{\varepsilon}{\sigma}\right)$

3.5.Định lý giới hạn trung tâm

Giả sử **có một số lượng lớn** các biến ngẫu nhiên $X_i\left(i = \overline{1,n}\right)$ phân phối **độc lập theo cùng một quy luật nào đó** với trung bình là m và phương sai là σ^2, thì tổng của chúng, $Y = \sum X_i$ $\left(X_i << Y\right)$, sẽ phân phối xấp xỉ quy luật chuẩn, với các tham số:

$E(Y) = \sum E(X_i) = n.m$ và $V(Y) = \sum V(X_i) = n\sigma^2$

***3.6.*Điều kiện hội tụ về quy luật chuẩn của quy luật Nhị Thức**

❖ Biến ngẫu nhiên $X \sim B(n;p)$ sẽ phân phối xấp xỉ quy luật chuẩn nếu **thỏa mãn 1 trong 2** điều kiện sau:

(1) $\begin{cases} n > 5 \\ \left|\sqrt{\dfrac{p}{1-p}} - \sqrt{\dfrac{1-p}{p}}\right| . \dfrac{1}{\sqrt{n}} < 0,3 \end{cases}$

(2) Nếu n lớn $(n \geq 100)$_theo ***định lý giới hạn trung tâm***. Ta đã biết, biến ngẫu nhiên phân phối *B(n;p)* là tổng của n biến ngẫu nhiên phân phối *A(p)*

❖ ***Lúc đó, ta sử dụng các công thức sau để tính xác suất:***

✓ $P(X = x) = \dfrac{1}{\sqrt{np(1-p)}}\Phi_0\left(\dfrac{x-np}{\sqrt{np(1-p)}}\right)$

✓ $P(x \leq X \leq x+m) = \Phi_0\left(\dfrac{x+m-np}{\sqrt{np(1-p)}}\right) - \Phi_0\left(\dfrac{x-np}{\sqrt{np(1-p)}}\right)$

4. Phân phối Student, phân phối Khi-Bình phương và phân phối Fisher

❖ **Giá trị tới hạn**

Giả sử có biến ngẫu nhiên X tuân theo quy luật G nào đó**. Giá trị tới hạn mức** α của quy luật G, kí hiệu x_α, là giá trị thỏa mãn: $P(X > x_\alpha) = \alpha$

(1) Khi-Bình phương: $\chi_\alpha^{2(n)}$ được xác định bằng cách tra bảng phụ lục. Trong đó n là số bậc tự do, α là mức giá trị tới hạn.

Ví dụ: tra giá trị $\chi_{0,95}^{2(6)}$, nhìn vào cột $\alpha = 0,95$ và dòng $n = 6$. Do không có dòng n = 6 nên chúng ta sẽ dùng dòng gần nhất với nó là $n = 5$. Ta được $\chi_{0,95}^{2(6)} \approx \chi_{0,95}^{2(5)} = 1,145$.

Khi-bình phương: $\chi_\alpha^2(n)$

n \ α	*0.975*	*0.95*	*0.05*	*0.025*
1	0.001	0.004	3.841	5.024
2	0.051	0.103	5.991	7.378
3	0.216	0.352	7.815	9.348
4	0.484	0.711	9.488	11.14
5	0.831	1.145	11.07	12.83
10	3.247	3.940	18.31	20.48

(2) Student: $t_\alpha^{(n)}$ là giá trị tới hạn student n bậc tự do mức α .

Tính chất: $t_{1-\alpha}^{(n)} = -t_\alpha^{(n)}$. Cách tra hoàn toàn tương tự như tra $\chi_\alpha^{2(n)}$

Ví dụ cần tra $t_{0,95}^{(9)}$, nhìn vào bảng ta không thấy cột $\alpha = 0,95$ nên ta sẽ dùng công thức phía trên để tra: $t_{0,95}^{(9)} = -t_{1-0,95}^{(9)} = -t_{0,05}^{(9)}$. Tra bảng ta thấy $t_{0,05}^{(9)} \approx t_{0,05}^{(10)} = 1,812$ nên $t_{0,95}^{(9)} = -1,812$

Student: $t_\alpha(n)$

n \ α	*0.1*	*0.05*	*0.025*
10	1.372	1.812	2.228
11	1.363	1.796	2.201
12	1.356	1.782	2.179
13	1.350	1.771	2.160

(3) Fisher: $f_\alpha^{(n_1;n_2)}$ là giá trị tới hạn fisher n_1, n_2 bậc tự do mức α .

Tính chất: $f_\alpha^{(n_1;n_2)} = \dfrac{1}{f_{1-\alpha}^{(n_2;n_1)}}$.

Cách tra: về cơ bản là cũng giống như cách tra các giá trị tới hạn phía trên.

Ví dụ: cần tra $f_{0,95}^{(25;40)} \approx f_{0,95}^{(24;39)} = \dfrac{1}{f_{0,05}^{(39;24)}} = \dfrac{1}{1,90} = 0,53$

Fisher: $f_\alpha(n_1, n_2)$

n_2	n_1 / α	**24**	**39**	**59**
15	*0.025*	2.70	2.59	2.53
	0.05	2.29	2.21	2.16
24	*0.025*	2.27	2.15	2.08
	0.05	1.98	1.90	1.84
39	*0.025*	2.02	1.89	1.82
	0.05	1.80	1.70	1.65
	[illegible]	[illegible]	[illegible]	[illegible]

5. Gợi ý các dạng toán thường gặp

5.1.Dạng 1: Các bài toán thuần phân phối Nhị Thức

(1) Chứng mình biến ngẫu nhiên phân phối Nhị Thức
(2) Tìm trung bình, phương sai, mốt
(3) Tính cách giá trị xác suất

5.2.Dạng 2: Các bài toán thuần phân phối chuẩn

(1) Tìm các tham số μ và σ^2 dựa vào xác giá trị xác suất
(2) Tìm các giá trị xác suất dựa vào công thức

5.3.Dạng 3: Các bài tổng hợp

(1) Phân phối Nhị thức xấp xỉ về quy luật chuẩn
(2) Phân phối chuẩn kết hợp Nhị thức-xấp xỉ về quy luật chuẩn
(3) Phân phối chuẩn kết hợp công thức xác suất đầy đủ, công thức Bernoulli
(4) Phân phối chuẩn với bài toán ra quyết định kinh tế

IV. Chương IV: Biến ngẫu nhiên 2 chiều và quy luật phân phối xác suất

1. Biến ngẫu nhiên 2 chiều

1.1.Định nghĩa

❖ Biến ngẫu nhiên 2 chiều $(X;Y)$ là một hệ hai biến ngẫu nhiên, bao gồm ngẫu biến X và ngẫu biến Y được xem xét đồng thời (cùng một lúc).

1.2.Phân loại:

❖ Nếu X và Y cùng rời rạc thì $(X;Y)$ là ngẫu biến 2 chiều rời rạc.

1.3.Các tham số đặc trưng <u>mới</u>

(1) ***Hiệp phương sai***: $Cov(X;Y) = E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] = E(XY) - E(X)E(Y)$

❖ Là tham số đo sự tương quan của 2 biến ngẫu nhiên X và Y.
Nếu $Cov(X;Y) > 0$ ta nói X, Y tương quan thuận chiều và ngược lại
Nếu $Cov(X;Y) = 0$ thì ta nói X và Y không tương quan nhau

❖ ***<u>Lưu ý</u>***: Nếu X độc lập Y thì $Cov(X;Y) = 0$ và do đó *X* không tương quan *Y*. Nhưng điều ngược lại thì không chắc xảy ra: *X* không tương quan *Y* thì chưa chắc *X* và *Y* độc lập nhau.

❖ ***Với X, Y bất kì:*** $V(aX + bY) = a^2V(X) + b^2(Y) + 2abCov(X;Y)$

(2) Hệ số tương quan: $\rho_{x,y} = \dfrac{Cov(X;Y)}{\sigma_x \sigma_y}$ đo mức độ tương quan tuyến tính giữa *X* và *Y*

Tính chất: $-1 \le \rho_{x,y} \le 1$

Nếu $\rho_{x,y} = \pm 1$ thì X và Y tương quan tuyến tính

Nếu $\rho_{x,y} = 0$ thì X và Y không tương quan

Nếu $\rho_{x,y}$ càng gần về ± 1 thì X và Y tương quan càng chặt chẽ và ngược lại, nếu $\rho_{x,y}$ càng gần về 0 thì X và Y tương quan càng yếu

2. Bảng phân phối xác suất đồng thời

2.1.Cấu tạo bảng phân phối

X \ Y	x_1	x_2	$\dots$	x_j	$\dots$	x_n
y_1	p_{11}	p_{12}	$\dots$	p_{1j}	$\dots$	p_{1n}
y_2	p_{21}	p_{22}	$\dots$	p_{2j}	$\dots$	p_{2n}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_i	p_{i1}	p_{i2}	$\dots$	p_{ij}	$\dots$	p_{in}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_m	p_{m1}	p_{m2}	$\dots$	p_{mj}	$\dots$	p_{mn}

Trong đó: x_j là các giá trị có thể có của X, y_i là các giá trị có thể có của Y. Các giá trị này được xếp tăng dần từ trái qua phải, từ trên xuống dưới.

$$p_{ij} = P\left(X = x_j; Y = y_i\right) \text{ và } \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} = p_{11} + p_{12} + .. + p_{ij} + .. + p_{nm} = 1$$

2.2.Một số biến đổi quan trọng

(1) $p_j = P(X = x_j) = \sum_{i=1}^{n} P\left(X = x_j; Y = y_i\right) =$

$= P\left(X = x_j; Y = y_1\right) + P\left(X = x_j; Y = y_2\right) + ... + P\left(X = x_j; Y = y_n\right)$

(2) $P\left(X = x_j / Y = y_i\right) = \dfrac{P\left(X = x_j; Y = y_i\right)}{P\left(Y = y_i\right)} = \dfrac{p_{ij}}{p_i}$

(3) $P\left(X + Y > x_j + y_i\right) = P\left(X = x_{j+1}; Y = y_{i+1}\right) + P\left(X = x_{j+1}; Y = y_{i+2}\right) + ... + P\left(X = x_{j+2}; Y = y_{i+1}\right) + ... +$

$+ P\left(X = x_{j+k}; Y = y_{i+k}\right) + ... + P\left(X = x_m; Y = y_n\right)$

(4) $E(XY) = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(p_{ij}.x_j y_i\right) = p_{11}.x_1 y_1 + p_{12}.x_2 y_1 + ... + p_{1m}.x_m y_1 + ... + p_{21}.x_1 y_2 + ... + p_{nm}.x_m y_n$

3. Bảng phân phối xác suất biên

***3.1.*Cách lập bảng phân phối xác suất biên**

❖ Bảng phân phối xác suất biên theo biến X có dạng:

X	x_1	x_2	x_3	$\dots$	x_m
Xác suất	p_1	p_2	p_3	$\dots$	p_m

Trong đó, p_j được xác định từ biến đổi số (1) ở phần **2.2**

❖ Hoàn toàn tương tự đối với bảng phân phối xác suất viên theo biến Y.

***3.2.*Các tham số đặc trưng**

❖ E(X), V(X), E(Y), V(Y) được tính toán như phần biến ngẫu nhiên 1 chiều.

4. Phân phối xác suất có điều kiện

❖ Bảng phân phối xác suất của biến X trong điều kiện $Y = y_i$ là:

X	x_1	x_2	x_3	$\dots$	x_m

Xác suất	p_1	p_2	p_3	$\cdots$	p_m

Trong đó, các p_j được xác định dựa theo biến đổi số (2) ở phần **2.2.**

Từ bảng này chúng ta tính được các ***kì vọng có điều kiện*** $E\left(X/Y=y_i\right)$ và ***phương sai có điều kiện*** $V\left(X/Y=y_i\right)$

❖ Hoàn toàn tương tự đối với bảng phân phối của biến Y trong điều kiện $X=x_j$

5. Hàm các biến ngẫu nhiên

5.1.Khái niệm*:* $Y=f(X_1,X_2,...,X_i)$ là một biến ngẫu nhiên được tạo bởi các biến ngẫu nhiên X_i cho trước thông qua quy tắc f.

Ví dụ: $Y=2X+X^2-3$; $Y=2X_1+3X_2+5X_3+1994$

5.2.Một số kết luận quan trọng:

❖ Nếu X_1,X_2,X_3 độc lập và cùng phân phối $N\left(\mu;\sigma^2\right)$ thì $Y=2X_1+3X_2+5X_3+1994$ cũng sẽ phân phối chuẩn với:

✓ $E(Y)=E\left(2X_1+3X_2+5X_3+1994\right)=2E(X_1)+3E\left(X_2\right)+5E\left(X_3\right)+1994=10\mu+1994$

✓ $V(Y)=V\left(2X_1+3X_2+5X_3+1994\right)=2^2V(X_1)+3^2V\left(X_2\right)+5^2V\left(X_3\right)=38\sigma^2$

❖ Một tổ hợp tuyến tính của các biến ngẫu nhiên độc lập phân phối chuẩn sẽ phân phối chuẩn.

❖ Nếu $X_1\sim B\left(n_1;p\right)$độc lập với $X_2\sim B\left(n_2;p\right)$ thì $X=X_1+X_2\sim B\left(n_1+n_2;p\right)$

6. Gợi ý các dạng toán thường gặp

6.1.*Bài toán lập bảng (tính các xác suất đồng thời, xác suất có điều kiện)*

6.2.*Tính toán các tham số đặc trưng; xét sự tương quan, độc lập, phụ thuộc giữa 2 ngẫu biến*

6.3.*Bài toán về hàm các biến ngẫu nhiên kết hợp với bảng phân phối đồng thời*