ĐẠI HỌC QUỐC GIA HÀ NỘI

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

KHOA TOÁN - CƠ - TIN HỌC

Xác suất

Probability

Mã lớp học phần: **MAT2405**

Sinh viên: **LƯU VĂN VIỆT**

Lớp: **A2K65 TOÁN - TIN**

Hà Nội, tháng 1 năm 2023

Mục lục

Chương 1

Biến cố và xác suất của biến cố

§1. Các quy tắc tính xác suất

1 Quy tắc cộng xác suất

Nếu A và B là hai biến cố xung khắc thì:

$$P(A \cup B) = P(A) + P(B)$$

Nếu A và B là hai biến cố bất kì thì:

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

2 Quy tắc chuyển sang biến cố đối

Biến cố đối của biến cố A, kí hiệu là $\overline{A}$. Xác suất của biến cố $\overline{A}$ được tính theo công thức:

$$P(\overline{A}) = 1 - P(A)$$

3 Quy tắc nhân

Nếu A và B là hai biến cố độc lập thì:

$$P(AB) = P(A).P(B)$$

Nếu $A_1, A_2, ..., A_n$ là các biến cố độc lập thì:

$$P(A_1A_2A_3...A_n) = P(A_1)P(A_2)...P(A_n)$$

§2. Xác suất có điều kiện

Giả sử A là một biến cố, B là một biến cố khác. Xác suất của B được tính trong điều kiện biết rằng A đã xảy ra được gọi là xác suất của B với điều kiện A, kí hiệu là $P(B|A)$ và được xác định bởi công thức:

$$P(B|A) = \frac{P(AB)}{P(B)}$$

Từ công thức trên, với A và B là hai biến cố bất kì, ta có:

$$P(AB) = P(B|A)P(A) = P(A|B)P(B)$$

§3. Công thức xác suất đầy đủ và công thức Bayes

1 Công thức xác suất đầy đủ

Các biến cố $B_1, B_2, ..., B_n$ được gọi là hệ đầy đủ các biến cố nếu chúng đôi một xung khắc với nhau $\left(P(B_iB_j) = 0, \forall i \neq j\right)$ và hợp của chúng là biến cố chắc chắn:

$$P(B_1 \cup B_2 \cup ... \cup B_n) = P(B_1) + P(B_2) + ... + P(B_n) = 1$$

Nếu $B_1, B_2, ..., B_n$ là hệ đầy đủ các biến cố thì với mỗi biến cố A, ta có:

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + ... + P(A|B_n)P(B_n)$$

2 Công thức Bayes

Nếu $B_1, B_2, .., B_n$ là hệ đầy đủ các biến cố và A là một biến cố với $P(A) > 0$ thì khi đó với mỗi $k = 1, 2, ..., n$:

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum\limits_{i=1}^{n} P(A|B_i)P(B_i)}$$

Chương 2

Đại lượng ngẫu nhiên rời rạc

§1. Phân bố nhị thức

ĐLNN X được gọi là có phân bố nhị thức với tham số n, p và kí hiệu là $X \sim B(n, p)$, ở đó $n \in \mathbb{N}$ và $0 < p < 1$ nếu $X(\Omega) = \{0, 1, 2, ..., n\}$ và:

$$P(X = k) = C_n^k p^k (1-p)^{n-k}$$

Ý nghĩa: Phép thử $\mathcal{T}$ được tiến hành n lần một cách độc lập với xác suất xuất hiện biến cố A là p. Gọi X là số lần xuất hiện biến cố A trong n lần thực hiện phép thử này thì khi đó X có phân bố nhị thức với tham số n và p.

Ví dụ: Xác suất làm ra một sản phẩm lỗi là $0,1$. Nhặt ra 1000 sản phẩm 1 cách độc lập và gọi X là số sản phẩm lỗi trong 1000 sản phẩm này. Khi đó X có phân bố nhị thức với tham số $n = 1000, p = 0,1$.

Một số đặc trưng của X:

- Kì vọng: $EX = np$
- Phương sai: $DX = np(1-p)$
- Mod: $Mod(X) = \lfloor (n+1)p \rfloor$

§2. Phân bố Poisson

ĐLNN X được gọi là có phân bố Poisson với tham số $\lambda > 0$ và kí hiệu là $X \sim P(\lambda)$ nếu $X(\omega) = \{0, 1, 2, ...\}$ và:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Ý nghĩa: Giả sử trong 1 đơn vị thời gian (giây, phút, giờ, năm) biến cố E xuất hiện trung bình c lần. Gọi X là số lần xuất hiện trong khoảng thời gian Δt thì khi đó, X có phân bố Poisson với tham số $\lambda = c\Delta t$.

Ví dụ: Ở một tổng đài bưu điện, các cú điện thoại gọi đến xuất hiện ngẫu nhiên, độc lập với nhau và tốc độ trung bình 2 cuộc gọi trong một phút. Khi đó nếu gọi X là số cú điện thoại xuất hiện trong hai phút thì X có phân bố Poisson với tham số $\lambda = c\Delta t = 2.2 = 4$.

Một số đặc trưng của X:

- Kì vọng: $EX = \lambda$
- Phương sai: $DX = \lambda$
- Mod: $Mod(X) = \lfloor \lambda \rfloor$

Chương 3

Luật số lớn và các định lý giới hạn

§1. Xấp xỉ phân bố nhị thức bằng phân bố Poisson

Cho (X_n) là dãy các ĐLNN có phân bố nhị thức, ở đó với mỗi n, X_n có phân bố nhị thức với tham số (n, p_n). Nếu n khá lớn và p_n khá bé thì phân bố nhị thức với tham số (n, p_n) có thể được xấp xỉ bởi phân bố Poisson với tham số $\lambda = np_n$. Xấp xỉ tốt nhất khi $n > 50, p_n < 0,1$.

Ví dụ: Xác suất làm ra 1 sản phẩm lỗi là 0,04. Chọn ngẫu nhiên 1000 sản phẩm và gọi X là số sản phẩm lỗi trong 1000 sản phẩm này. Tính xác suất để có đúng 50 sản phẩm lỗi.

Do X là số sản phẩm lỗi trong 1000 sản phẩm nên $X \sim B(1000; 0,04)$.

$\Rightarrow X$ có phân bố xấp xỉ phân bố Poisson với tham số $\lambda = 1000.0,04 = 40$.

Gọi $\widetilde{X} \sim P(40) \Rightarrow P(X = 50) \approx P(\widetilde{X} = 50) = e^{-40}\dfrac{40^{50}}{50!} = 0,0177$

§2. Xấp xỉ phân bố nhị thức bằng phân bố chuẩn

Phân bố nhị thức $B(n, p)$ có thể được xấp xỉ bởi phân bố chuẩn:

$$N(np, np(1-p))$$

Xấp xỉ tốt nhất khi $np > 5, n(1-p) > 5$ hoặc $np(1-p) > 20$

Giả sử $X \sim B(n, p)$ và có thể được xấp xỉ bởi phân bố chuẩn $\widetilde{X} \sim N(np, np(1-p))$. Khi đó ta đã xấp xỉ phân bố rời rạc bởi phân bố liên tục nên ta cần hiệu chỉnh để giảm sai số:

$$P(X < k) \approx P(\widetilde{X} < k - 0,5)$$

$$P(X \leq k) \approx P(\widetilde{X} < k + 0,5)$$

$$P(k_1 < X < k_2) \approx P(k_1 + 0,5 < \widetilde{X} < k_2 - 0,5)$$

$$P(k_1 \leq X < k_2) \approx P(k_1 - 0,5 < \widetilde{X} < k_2 - 0,5)$$

$$P(k_1 < X \leq k_2) \approx P(k_1 + 0,5 < \widetilde{X} < k_2 + 0,5)$$

$$P(k_1 \leq X \leq k_2) \approx P(k_1 - 0,5 < \widetilde{X} < k_2 + 0,5)$$

Ví dụ: Một kí túc xá có 650 sinh viên. Xác suất để một sinh viên đến xem phim tại câu lạc bộ vào tối thứ 7 là $0,7$.

a. Tính xác suất để số sinh viên đi xem phim vào tối thứ 7 ít hơn 440.

b. Cần chuẩn bị bao nhiêu ghế để với xác suất 0,99 ta có thể đảm bảo đủ ghế cho người đến xem.

Lời giải

a.

Gọi X là số sinh viên đến xem phim thì $X \sim B(650; 0,7)$

Do $650.0,7.0,3 = 136,5 > 20$ nên ta có thể xấp xỉ X bởi phân bố chuẩn với tham số $\mu = 650.0,7 = 455$ và $\sigma^2 = 650.0,7.0,3 = 136,5$.

Gọi $\widetilde{X} \sim N(455; 136,5)$

Ta có: $P(X < 440) \approx P(\widetilde{X} < 439,5) = \phi(\dfrac{439,5 - 455}{\sqrt{136,5}}) = \phi(-1,327) = 1 - \phi(1,327) = 0,0923$

b.

Giả sử k là số ghế cần chuẩn bị thì k phải thỏa mãn: $P(X > k) \leq 0,01$

$\Rightarrow P(X > k) \approx P(\widetilde{X} > k + 0,5) \leq 0,01$

$\Rightarrow P(\widetilde{X} < k + 0,5) \geq 0,99 = \phi(2,326)$

$\Rightarrow \dfrac{k + 0,5 - 455}{\sqrt{135,5}} > 2,326 \Leftrightarrow k \leq 481,7$

§3. Bài tập

Bài 1: Trong một nhà máy có ba phân xưởng A, B, C tương ứng làm ra $25\%, 30\%, 45\%$ tổng số sản phẩm của nhà máy. Biết rằng xác suất làm ra một sản phẩm hỏng của phân xưởng A, B, C lần lượt là $0,03; 0,015; 0,03$. Chọn ngẫu nhiên một sản phẩm của nhà máy.

a. Tính xác suất để sản phẩm được chọn là sản phẩm hỏng.

b. Biết rằng sản phẩm được chọn là sản phẩm tốt. Tính xác suất để sản phẩm đó lấy từ phân xưởng B.

c. Giả sử lấy 1000 sản phẩm của phân xưởng một cách độc lập. Tính xác suất để có đúng 50 sản phẩm hỏng.

d. Giả sử lấy 1000 sản phẩm của phân xưởng một cách độc lập. Tính xác suất để số sản phẩm hỏng nằm trong khoảng $(15, 50]$.

Bài 2: Trong một nhà máy có 3 loại máy A, B, C với tỉ lệ là $0,25; 0,45; 0,3$. Chọn ngẫu nhiên một máy, từ máy đó sản xuất 20 sản phẩm, 20 sản phẩm vừa được sản xuất sẽ đóng thành một lô hàng. Một lô hàng được đánh giá là tốt nếu số sản phẩm lỗi nhiều nhất là một sản phẩm. Biết rằng xác suất để làm ra một sản phẩm lỗi của các máy A, B, C lần lượt là $0,2; 0,2; 0,3$

a. Tính xác suất để có được một lô hàng tốt

b. Chọn ngẫu nhiên một lô hàng thì kiểm tra thấy lô hàng đó là lô hàng tốt. Tính xác suất để lô hàng đó được sản xuất từ máy A.

c. Chọn ngẫu nhiên 15000 lô hàng, số lô hàng tốt tin chắc nhất là bao nhiêu. **Bài 3:** Có hai lớp học A, B. Lớp A có 5 nữ và 5 nam; lớp B có 3 nữ và 6 nam. Người ta điều chuyển 1 em từ lớp A sang lớp B. Nhà trường gọi một em lớp B lên gặp. Tính xác suất để:

a. Sinh viên được gọi lên là nữ

b. Biết rằng sinh viên được gọi lên là nữ, tìm xác suất để sinh viên chuyển từ lớp A sang lớp B là nữ.

Bài 4: Trong một cái túi có 6 đồng xu thật và 4 đồng xu giả. Biết rằng các đồng xu thật đều cân đối đồng chất còn các đồng xu giả thì không; và nếu ta tung 1 đồng xu giả thì xác suất được mặt sấp sẽ là $0,3$. Ta làm phép thử như sau: rút hú họa một đồng xu từ túi rồi tung đồng xu đó.

a. Tính xác suất để ta được mặt sấp

b. Giả sử ta tung đồng xu được mặt ngửa. Tìm xác suất để đồng xu đó là giả.

c. Giả sử có 100 người làm phép thử trên một các độc lập với nhau. Tính xác suất để có đúng 35 người được mặt sấp.

d. Tính xác suất để có không quá 35 người tung được mặt sấp.

e. Hỏi số người tung được mặt sấp tin chắc nhất là bao nhiêu?

Bài 5: Máy bay có thể xuất hiện một cách ngẫu nhiên trong ba vị trí A, B, C. Để bắn rơi máy bay, người ta bố trí 2 khẩu pháo đặt tại A, 1 khẩu pháo đặt tại B và 1 khẩu pháo đặt tại C. Biết rằng xác suất bắn trúng máy bay của mỗi khẩu pháo là $0,7$ và các khẩu pháo hoạt động động lập với nhau.

a. Tính xác suất để máy bay bị bắn trúng.

b. Giả sử máy bay không bị bắn trúng. Tính xác suất để máy bay xuất hiện ở A.

c. Giả sử xuất hiện 100 máy bay. Tính xác suất để có đúng 70 máy bay bị bắn trúng.

d. Tính xác suất xuất để có không quá 70 máy bay bị bắn trúng.

Bài 6: Ở một nông trại chăn nuôi lợn người ta phát hiện lợn bị mắc một loại bệnh mà chỉ do virus A và B gây ra, người ta biết rằng khi lợn bị bệnh này thì 50% là trong cơ thể chỉ có virus A, 80% trong cơ thể chỉ có virus B. Hơn nữa, nếu trong cơ thể chỉ có virus A thì xác suất để chữa khỏi là $0,8$, nếu trong cơ thể chỉ có virus B thì xác suất để chữa khỏi là $0,6$ và nếu có cả hai loại virus thì xác suất để chữa khỏi là $0,4$.

a. Tính xác suất để một con lợn bị bệnh được chữa khỏi.

b. Biết nông trại có 500 con lợn bị bệnh, gọi X là số lợn được chữa khỏi. Khi đó X có phân bố gì? Tính EX?

Bài 7: Một xạ thủ bắn lần lượt 5 viên đạn vào 1 mục tiêu. Xác suất để bắn trúng mục tiêu của mỗi viên đạn $0,8$ và các lần bắn là độc lập với nhau.

a. Tính xác suất để xạ thủ bắn trúng 4 viên, biết rằng không có quá 4 viên đạn trúng mục tiêu.

b. Hỏi xạ thủ phải có ít nhất bao nhiêu viên đạn để xác suất của biến cố: "Bắn trúng mục tiêu" không nhỏ hơn $0,95$.

Chương 4

Đại lượng ngẫu nhiên liên tục

§1. Hàm mật độ xác suất, hàm phân bố xác suất và một số đặc trưng

1 Hàm mật độ xác suất

Hàm số $f(x)$ xác định trên toàn trục số được gọi là hàm mật độ của ĐLNN liên tục X nếu:

(i) $f(x) \leq 0, \forall x \in \mathbb{R}$

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

(iii) $P(a < X < b) = \displaystyle\int_a^b f(x)dx$

2 Hàm phân bố xác suất

Hàm phân bố xác suất của ĐLNN X, kí hiệu bởi $F(x)$, là hàm xác định bởi mọi số thực x theo công thức sau:

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt$$

3 Một số đặc trưng

- Kì vọng: $EX = \displaystyle\int_{-\infty}^{+\infty} xf(x)dx$
- Phương sai: $DX = EX^2 - (EX)^2 = \displaystyle\int_{-\infty}^{+\infty} x^2 f(x)dx - (EX)^2$
- Mod: $Mod(X) = m \Leftrightarrow f(m) = \max f(x)$
- Median: $Med(X) = m \Leftrightarrow P(X < m) = P(X > m) \Leftrightarrow F(m) = \dfrac{1}{2}$

4 Hàm của một ĐLNN

Giả sử X là một ĐLNN và $g(x)$ là một hàm đã cho trước. Xét ĐLNN Y xác định bởi:

$$Y = g(X)$$

Khi đó Y là hàm của ĐLNN X và kì vọng của Y được tính bởi công thức:

$$EY = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

§2. Một số phân bố quan trọng

1 Phân bố chuẩn

ĐLNN Z được gọi là có phân bố chuẩn tắc nếu hàm mật độ của nó là:

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

ĐLNN X được gọi là có phân bố chuẩn với tham số μ và σ^2 nếu ĐLNN $Z = \dfrac{X-\mu}{\sigma}$ có phân bố chuẩn tắc. Khi đó, ta kí hiệu:

$$X \sim N(\mu, \sigma^2)$$

Các đặc trưng:

- Kì vọng: $EX = \mu$
- Phương sai: $DX = \sigma^2$
- Mod: $Mod(X) = \mu$
- Median: $Med(X) = \mu$

Tính các xác suất liên quan tới X:

- $P(X < a) = \phi(\dfrac{a-\mu}{\sigma})$
- $P(X > a) = 1 - P(X < a) = 1 - \phi(\dfrac{a-\mu}{\sigma})$
- $P(a < X < b) = \phi(\dfrac{b-\mu}{\sigma}) - \phi(\dfrac{a-\mu}{\sigma})$

§3. Phân bố mũ

ĐLNN X được gọi là có phân bố mũ với tham số $\lambda > 0$ nếu nó có hàm mật độ như sau:

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

Ý nghĩa: Trong một số giả thiết cho trước, khoảng thời gian giữa 2 lần xuất hiện của một biến cố E sẽ có phân bố mũ.

Các đặc trưng:

- Xác suất: $F(x) = P(X < x) = 1 - e^{-\lambda x}$
- Kì vọng: $EX = \dfrac{1}{\lambda}$
- Phương sai: $DX = \dfrac{1}{\lambda^2}$

Ví dụ: Giả sử tuổi thọ trung bình của một mạch điện tử trong máy tính là một ĐLNN có phân bố mũ với kì 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 này bán ra phải thay thế trong thời gian bảo hành?

Lời giải

Ta có: $\lambda = \dfrac{1}{EX} = \dfrac{1}{6,25} \Rightarrow X$ có phân bố mũ với tham số $\lambda = \dfrac{1}{6,25}$
Ta có: $P(X \leq 5) = 1 - e^{-5.\lambda} = 1 - e^{-0,8} = 0,5506$

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

1 Phân bố đều

ĐLNN X được gọi là có phân bố đều trên đoạn $[a, b]$ nếu X có thể nhận bất kì giá trị nào trên $[a, b]$ với xác suất như nhau và không nhận giá trị nào bên ngoài $[a, b]$:

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

Hàm phân bố của X được tính như sau:

$$F(x) = \begin{cases} 0 & \text{nếu } x < a \\ \dfrac{x-a}{b-a} & \text{nếu } a < x < b \\ 1 & \text{nếu } x > b \end{cases}$$

Các đặc trưng:

- Xác suất để $X \in [\alpha, \beta]$: $P(\alpha < X < \beta) = \dfrac{\beta - \alpha}{b - a}$
- Kì vọng: $EX = \dfrac{a+b}{2}$
- Phương sai: $DX = \dfrac{(b-a)^2}{12}$
- Mod: Mod của phân bố đều là bất cứ điểm nào trong $[a, b]$
- Median: $Med(X) = \dfrac{a+b}{2}$

§4. Bài tập

Bài 1: Trọng lượng X của các đơn hàng được gửi từ một bưu cục quận Thanh Xuân là một biến ngẫu nhiên liên tục có hàm mật độ:

$$f(x) = \begin{cases} \dfrac{70}{cx^2} & \text{nếu } 1 \le x \le 70 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm phương sai của biến ngẫu nhiên $Y = X^2 - 1$.

b. Nếu phí gửi tài liệu là 15000/kg thì phí trung bình của một đơn hàng được gửi từ bưu cục là bao nhiêu?

c. Tính xác suất của một đơn hàng có trọng lượng nhỏ hơn 35kg.

Bài 2: Biết rằng tuổi thọ (Tính theo năm) T của mỗi camera là biến ngẫu nhiên có hàm mật độ phân phối:

$$f(t) = \begin{cases} 0,5e^{-0,5t} & \text{nếu } t > 0 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm trung bình thời gian hoạt động của Camera.

b. Tìm xác suất để sau 2 năm camera còn hoạt động.

c. Để kiểm tra an ninh khu vực, người ta lắp đặt 3 camera cùng loại. Tìm xác suất để sau hai năm có ít nhất một camera hoạt động.

Bài 3: Một mạch điện của một nhà máy sản xuất có tuổi thọ (tính bằng năm) là một đại lượng ngẫu nhiên có phân bố mũ, biết tuổi thọ trung bình là 5 năm.

a. Nhà sản xuất muốn mạch điện bán ra không quá 20% sản phẩm phải bảo hành. Khi đó nhà sản xuất cần đưa ra thời gian bảo hành của mạch điện này tối đa là bao lâu?

b. Nếu một máy sản xuất có 2 loại mạch này mắc nối tiếp với nhau (máy chỉ hỏng nếu và chỉ nếu một trong hai mạch hỏng). Tính xác suất để thời gian máy tính hoạt động của máy là bé hơn 4 năm.

c. Một xí nghiệp mua 100 mạch điện loại này. Dùng định lý giới hạn trung tâm ước lượng xác suất để trong 100 mạch này có nhiều hơn 60 mạch bị hỏng trước 4 năm.

Bài 4: Một hệ thống linh kiện điện tử mắc nối tiếp gồm 3 linh kiện I, II, III. Nếu một linh kiện bị hỏng thì cả hệ thống sẽ bị hỏng. Biết tuổi thọ tính bằng năm của các linh kiện I, II và III tuân theo phân bố mũ với tham số lần lượt là $2; 2,5$ và 3.

a. Tìm xác suất để linh kiện I bị hỏng trong vòng 3 năm.

b. Tìm hàm mật độ của đại lượng ngẫu nhiên thể hiện tuổi thọ của hệ thống. Từ đó tìm tuổi thọ trung bình của hệ thống.

Chương 5

Đại lượng ngẫu nhiên liên tục nhiều chiều

§1. Hàm mật độ và hàm phân bố của ĐLNN liên tục nhiều chiều

1 Hàm mật độ

Nếu $f(x,y)$ là hàm mật độ đồng thời của X, Y thì $f(x,y)$ thỏa mãn:

(i) $f(x,y) \geq 0 \forall (x,y) \in \mathbb{R}^2$

(ii) $\displaystyle\iint_{\mathbb{R}^2} f(x,y)dxdy = 1$

(iii) $P((x,y) \in D) = \displaystyle\iint_{D} f(x,y)dxdy$

Hàm mật độ từng thành phần của X và Y được xác định như sau:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx$$

Kì vọng cặp ĐLNN (X, Y):

$$EXY = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy$$

2 Hàm phân bố

$F(x,y) = \displaystyle\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv$

$F_X(x) = \lim\limits_{y\to+\infty} F(x,y)$

$F_Y(y) = \lim\limits_{x\to+\infty} F(x,y)$

3 Tính độc lập của hai ĐLNN

Hai ĐLNN X và Y là độc lập nếu và chỉ nếu hàm mật độ đồng thời là tích của hai hàm mật độ của hai thành phần:

$$f(x,y) = f_X(x)f_Y(y)$$

4 Covarian và Hệ số tương quan

Covarian của hai ĐLNN X và Y là một số xác định bởi công thức:

$$cov(X,Y) = E((X - EX)(Y - EY))$$

Chúng ta thường sử dụng công thức sau khi tính toán Covarian trong thực tế:

$$cov(X,Y) = EXY - EX.EY = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy - EX.EY$$

Hệ số tương quan của X và Y, kí hiệu bởi $\rho(X,Y)$, được định nghĩa bởi công thức:

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X.\sigma_y}$$

§2. Phân bố có điều kiện và kì vọng có điều kiện

Xác suất $P(Y < y)$ tính trong điều kiện biết rằng $X = x$ xác định cho ta một hàm của y với tham số x (kí hiệu là $F(y|x)$). $F(y|x)$ được gọi là hàm phân bố của Y với điều kiện $X = x$

Giả sử $f(x,y)$ là hàm mật độ đồng thời của X, Y, còn $f_X(x)$ là hàm mật độ của X. Khi đó:

$$F(y|x) = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)}dv$$

Hàm mật độ có điều kiện của Y với điều kiện $X = x$ và ký hiệu là $f(y|x)$ được xác định bởi công thức:

$$f(y|x) = \frac{f(x,y)}{f_X(x)}$$

§3. Hàm của Đại lượng ngẫu nhiên hai chiều

Giả sử X và Y là hai ĐLNN liên tục với hàm mật độ đồng thời $f(x,y)$ và ký hiệu:

$$D = \{(x,y) : f(x,y) > 0\}$$

Giả sử $u(x,y)$ và $v(X,y)$ là hàm hai biến sao cho ánh xạ $T(x,y) = (u(x,y), v(x,y))$ thỏa mãn điều kiện $J(u,v) \neq 0$. Khi đó, cặp ĐLNN (U,V) xác định bởi:

$$U = u(X,Y), V = v(X,Y)$$

sẽ có hàm mật độ $g(u,v)$ như sau:

$$g(u,v) = \begin{cases} f(x(u,v), y(u,v)).|J(u,v)| & \text{nếu } (u,v) \in S \\ 0 & \text{nếu ngược lại} \end{cases}$$

§4. Bài tập

Bài 1: Cho X, Y là 2 đại lượng ngẫu nhiên liên tục có hàm mật độ:

$$f(x,y) = \begin{cases} kx^2y & \text{nếu } 0 < x, y < 2 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm k và tính xác suất để $X > 2Y$.

b. Tìm hàm mật độ từng thành phần của X và Y. Từ đó kiểm tra tính độc lập của X và Y.

c. Tìm hàm mật độ của biến ngẫu nhiên $Z = X^2 - Y$

Bài 2: Cho (X, Y) là một vector ngẫu nhiên 2 chiều thỏa mãn X có phân phối đều liên tục trên đoạn $[0, 10]$ và hàm mật độ có điều kiện của Y với điều kiện $X = x$ là:

$$f(y|x) = xe^{-xy} \text{ khi } y > 0$$

a. Tính kỳ vọng có điều kiện $E(Y|X = 2)$

b. Tìm hàm mật độ của biến ngẫu nhiên Y và tính phương sai của Y.

Bài 3: Cho hàm mật độ đồng thời của (X, Y) là:

$$f(x,y) = x + y \text{ khi } 0 \le x, y \le 1$$

Hãy tìm hệ số tương quan của X và Y.

Bài 4: Cho hàm mật độ xác suất đồng thời của vector ngẫu nhiên hai chiều được xác định bởi:

$$f(x,y) = \begin{cases} 12xy(1-x) & \text{nếu } 0 < x, y < 1 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm hàm mật độ xác suất của X và Y.

b. Kiểm tra xem X và Y có độc lập với nhau hay không?

c. Tính $P(Y < 0,5|X > 0,5)$

Bài 5: Cho hàm mật độ đồng thời của X và Y là:

$$f(x,y) = \begin{cases} \dfrac{e^{-x}e^{-\frac{x}{y}}}{y} & \text{nếu } x > 0, y > 0 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm hàm mật độ $f_Y(y)$ của Y.

b. Tính $E(X|Y = y)$

Bài 6: Biết hàm mật độ đồng thời của (X, Y) là:

$$f(x,y) = \begin{cases} ce^{-2y} & \text{nếu } 0 < x < y \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Xác định c?

b. Tìm hàm mật độ của Y

c. Tính $cov(X, Y)$

Chương 6

Đề thi các năm

Đề thi xác suất cuối kì I ngành Toán - Tin năm học 2021 - 2022

Câu 1: Trong một nhà máy có ba phân xưởng A, B, C tương ứng làm ra $25\%, 30\%, 45\%$ tổng số sản phẩm của nhà máy. Biết rằng xác suất làm ra một sản phẩm hỏng của phân xưởng A, B, C lần lượt là $0,03; 0,015; 0,03$. Chọn ngẫu nhiên một sản phẩm của nhà máy.

a. Tính xác suất để sản phẩm được chọn là sản phẩm hỏng.

Đáp án: $0,0255$

b. Biết rằng sản phẩm được chọn là sản phẩm tốt. Tính xác suất để sản phẩm đó lấy từ phân xưởng B.

Đáp án: $0,3032$

c. Giả sử lấy 1000 sản phẩm của phân xưởng một cách độc lập. Tính xác suất để có đúng 50 sản phẩm hỏng.

Đáp án: $5,88.10^{-6}$

d. Giả sử lấy 1000 sản phẩm của phân xưởng một cách độc lập. Tính xác suất để số sản phẩm hỏng nằm trong khoảng $(15, 50]$.

Đáp án: $0,9773$

Câu 2: Một xạ thủ bắn lần lượt 5 viên đạn vào 1 mục tiêu. Xác suất để bắn trúng mục tiêu của mỗi viên đạn $0,8$ và các lần bắn là độc lập với nhau.

a. Tính xác suất để xạ thủ bắn trúng 4 viên, biết rằng không có quá 4 viên đạn trúng mục tiêu.

Đáp án: $0,6092$

b. Hỏi xạ thủ phải có ít nhất bao nhiêu viên đạn để xác suất của biến cố: "Bắn trúng mục tiêu" không nhỏ hơn $0,95$.

Đáp án: 2

Câu 3: Cho X, Y là 2 đại lượng ngẫu nhiên liên tục có hàm mật độ:

$$f(x,y) = \begin{cases} kx^2y & \text{nếu } 0 < x, y < 2 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm k và tính xác suất để $X > 2Y$.

Đáp án: $k = \dfrac{3}{16}, P(X > 2Y) = \dfrac{3}{20}$

b. Tìm hàm mật độ từng thành phần của X và Y. Từ đó kiểm tra tính độc lập của X và Y.

Đáp án:

$$f_X(x) = \begin{cases} \dfrac{3}{8}x^2 & \text{nếu } 0 < x < 2 \\ 0 & \text{nếu ngược lại} \end{cases} \qquad f_Y(y) = \begin{cases} \dfrac{y}{2} & \text{nếu } 0 < y < 2 \\ 0 & \text{nếu ngược lại} \end{cases}$$

Do $f_X(x).f_Y(y) = f(x,y)$ nên X và Y độc lập.

c. Tìm hàm mật độ của biến ngẫu nhiên $Z = X^2 - Y$

Đáp án:

Đặt $U = X^2$, ta tính được hàm mật độ của U: $f_U(u) = \begin{cases} \dfrac{3\sqrt{u}}{16} & \text{nếu } 0 < u < 4 \\ 0 & \text{nếu ngược lại} \end{cases}$

Do X và Y độc lập nên U và Y độc lập, do đó hàm mật độ đồng thời của U và Y là:

$$f(u,y) = \begin{cases} \dfrac{3y\sqrt{u}}{32} & \text{nếu } 0 < u < 4, 0 < y < 2 \\ 0 & \text{nếu ngược lại} \end{cases}$$

Với $Z = X^2 - Y = U - Y$, đặt $\begin{cases} T = U \\ V = U - Y \end{cases} \Rightarrow \begin{cases} U = T \\ Y = T - V \end{cases} \Rightarrow |J| = 1 \neq 0$

$D = \{(u,y)|0 < u < 4, 0 < y < 2\} \Rightarrow S = \{(t,v)|0 < t < 4, t - 2 < v < t\}$

$$\Rightarrow g(t,v) = \begin{cases} f(u(t,v), y(t,v))|J| = \dfrac{3(t-v)\sqrt{t}}{32} & \text{nếu } (u,v) \in S \\ 0 & \text{nếu ngược lại} \end{cases}$$

$$g_V(v) = \begin{cases} \dfrac{3}{32}(\dfrac{2}{5}(v+2)^{\frac{5}{2}} - \dfrac{2}{3}(v+2)^{\frac{3}{2}}v) & \text{nếu } -2 < v < 0 \\ \dfrac{3}{32}(\dfrac{2}{5}(v+2)^{\frac{5}{2}} - \dfrac{2}{3}(v+2)^{\frac{3}{2}}v + \dfrac{4}{15}v^{\frac{5}{2}}) & \text{nếu } 0 < v < 2 \\ \dfrac{3}{32}(\dfrac{4}{15}v^{\frac{5}{2}} - \dfrac{16}{3}v + \dfrac{65}{4}) & \text{nếu } 2 < v < 4 \\ 0 & \text{nếu ngược lại} \end{cases}$$

Đề thi xác suất học kì I ngành Toán-Tin năm học 2020 - 2021

Câu 1: Cho X là biến ngẫu nhiên có phân bố $P(X=-1)=P(X=2)=0,5$ và Y là biến ngẫu nhiên có phân phối nhị thức với tham số $n=3$ và $p=0,6$. Giả sử rằng X và Y là độc lập.

a. Tìm bảng phân bố xác suất đồng thời của X và Y.

b. Tính xác suất $P(X+Y\geq 4)$.

Câu 2: Trọng lượng X của các đơn hàng được gửi từ một bưu cục quận Thanh Xuân là một biến ngẫu nhiên liên tục có hàm mật độ:

$$f(x)=\begin{cases}\dfrac{70}{cx^2} & \text{nếu } 1\leq x\leq 70\\ 0 & \text{nếu ngược lại}\end{cases}$$

a. Tìm phương sai của biến ngẫu nhiên $Y=X^2-1$.

b. Nếu phí gửi tài liệu là 15000/kg thì phí trung bình của một đơn hàng được gửi từ bưu cục là bao nhiêu?

c. Tính xác suất của một đơn hàng có trọng lượng nhỏ hơn 35kg.

Câu 3: Cho (X,Y) là một vector ngẫu nhiên 2 chiều thỏa mãn X có phân phối đều liên tục trên đoạn $[0,10]$ và hàm mật độ có điều kiện của Y với điều kiện $X=x$ là:

$$f(y|x)=xe^{-xy} \text{ khi } y>0$$

a. Tính kỳ vọng có điều kiện $E(Y|X=2)$

b. Tìm hàm mật độ của biến ngẫu nhiên Y và tính phương sai của Y.

Câu 4: Cho hàm mật độ đồng thời của (X,Y) là:

$$f(x,y)=x+y \text{ khi } 0\leq x,y\leq 1$$

Hãy tìm hệ số tương quan của X và Y.

Đề thi xác suất học kì I năm học 2019 - 2020

Câu 1: Một công ti bảo hiểm chia người dân thành hai nhóm: ít rủi ro và rủi ro. Số liệu thu thập được cho thấy mỗi năm một người thuộc nhóm ít rủi ro và rủi ro có số lần gặp tai nạn là phân bố Poisson với tham số λ tương ứng lần lượt là $0,1$ và $0,2$. Cho biết 70% người dân thuộc nhóm ít rủi ro, 30% người dân thuộc nhóm rủi ro.

a. Một người thuộc nhóm ít rủi ro thì xác suất gặp tai nạn trong một năm là bao nhiêu?

b. Một người thuộc nhóm rủi ro thì xác suất gặp ít nhất 2 tai nạn trong 3 năm là bao nhiêu?

c. Chọn ngẫu nhiên một người. Xác suất để người đó gặp tai nạn trong một năm là bao nhiêu?

Câu 2: Biết hàm mật độ đồng thời của (X, Y) là:

$$f(x,y) = \begin{cases} ce^{-2y} & \text{nếu } 0 < x < y \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Xác định c?

b. Tìm hàm mật độ của Y

c. Tính $cov(X, Y)$

Câu 3: Một mạch điện của một nhà máy sản xuất có tuổi thọ (tính bằng năm) là một đại lượng ngẫu nhiên có phân bố mũ, biết tuổi thọ trung bình là 5 năm.

a. Nhà sản xuất muốn mạch điện bán ra không quá 20% sản phẩm phải bảo hành. Khi đó nhà sản xuất cần đưa ra thời gian bảo hành của mạch điện này tối đa là bao lâu?

b. Nếu một máy sản xuất có 2 loại mạch này mắc nối tiếp với nhau (máy chỉ hỏng nếu và chỉ nếu một trong hai mạch hỏng). Tính xác suất để thời gian máy tính hoạt động của máy là bé hơn 4 năm.

c. Một xí nghiệp mua 100 mạch điện loại này. Dùng định lý giới hạn trung tâm ước lượng xác suất để trong 100 mạch này có nhiều hơn 60 mạch bị hỏng trước 4 năm.

Đề thi Xác suất ngành Toán - Tin năm học 2018-2019

Câu 1: Một công ty bảo hiểm chia người dân thành ba nhóm: ít rủi ro, rủi ro trung bình và rủi ro. Số liệu thu thập được cho thấy trong 3 năm một người thuộc nhóm ít rủi ro, rủi ro trung bình và rủi ro có số lần gặp tai nạn là phân bố Poisson với tham số λ tương ứng là $0,1; 0,2; 0,3$. Cho biết 50% người dân thuộc nhóm ít rủi ro, 30% người dân thuộc nhóm rủi ro trung bình và 20% người dân thuộc nhóm rủi ro.

a. Một người thuộc nhóm ít rủi ro thì xác suất gặp tai nạn trong ba năm là bao nhiêu?

b. Chọn ngẫu nhiên 1 người. Xác suất để người đó gặp tai nạn trong ba năm là bao nhiêu? Ước lượng số người gặp tai nạn trong 3 năm của một tỉnh biết số sân của tỉnh đó là 100000 dân.

Câu 2: Một hệ thống linh kiện điện tử mắc nối tiếp gồm 3 linh kiện I, II, III. Nếu một linh kiện bị hỏng thì cả hệ thống sẽ bị hỏng. Biết tuổi thọ tính bằng năm của các linh kiện I, II và III tuân theo phân bố mũ với tham số lần lượt là $2; 2,5$ và 3.

a. Tìm xác suất để linh kiện I bị hỏng trong vòng 3 năm.

b. Tìm hàm mật độ của đại lượng ngẫu nhiên thể hiện tuổi thọ của hệ thống. Từ đó tìm tuổi thọ trung bình của hệ thống.

Câu 3: Cho hàm mật độ đồng thời của X và Y là:

$$f(x,y) = \begin{cases} \dfrac{e^{-x}e^{-\frac{x}{y}}}{y} & \text{nếu } x > 0, y > 0 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm hàm mật độ $f_Y(y)$ của Y.

b. Tính $E(X|Y = y)$

Câu 4: Cho (X, Y) là vector ngẫu nhiên hai chiều có bảng phân bố như sau:

X \ Y	-1	0
2	$\frac{1}{4}$	$\frac{1}{8}$
3	$\frac{1}{8}$	$\frac{1}{2}$

Tìm $E(X|Y = 0)$ và kiểm tra tính độc lập của X và Y.

Đề thi Xác suất học kì I năm học 2017 - 2018

Câu 1: Cho X và Y là hai ĐLNN độc lập có cùng phân bố chuẩn tắc $N(0,1)$. Sử dụng hàm đặc trưng, chứng minh rằng biến ngẫu nhiên $Z = X + Y$ cũng có phân bố chuẩn. Tìm kì vọng và phương sai của Z.

Câu 2: Biết rằng tuổi thọ (Tính theo năm) T của mỗi camera là biến ngẫu nhiên có hàm mật độ phân phối:

$$f(t) = \begin{cases} 0,5e^{-0,5t} & \text{nếu } t > 0 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm trung bình thời gian hoạt động của Camera.

b. Tìm xác suất để sau 2 năm camera còn hoạt động.

c. Để kiểm tra an ninh khu vực, người ta lắp đặt 3 camera cùng loại. Tìm xác suất để sau hai năm có ít nhất một camera hoạt động.

Câu 3: Cho hàm mật độ xác suất đồng thời của vector ngẫu nhiên hai chiều được xác định bởi:

$$f(x,y) = \begin{cases} 12xy(1-x) & \text{nếu } 0 < x, y < 1 \\ 0 & \text{nếu ngược lại} \end{cases}$$

a. Tìm hàm mật độ xác suất của X và Y.

b. Kiểm tra xem X và Y có độc lập với nhau hay không?

c. Tính $P(Y < 0,5 | X > 0,5)$

Câu 4: CMR nếu $X_1, X_2, ..., X_n, ...$ là dãy BNN độc lập cùng phân phối với $EX_1 = 0$ và $DX_1 = 1$ thì:

$$\lim_{n\to\infty} P(\frac{S_n}{\sqrt{n}} < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}$$

trong đó $S_n = X_1 + X_2 + ... + X_n$

Câu 5:

a. Định nghĩa xác suất có điều kiện của biến cố A với điều kiện B xảy ra. Khi nào thì A và B độc lập.

b. Khái niệm họ đầy đủ các biến cố. Chứng minh công thức xác suất đầy đủ.

c. Chứng minh công thức Bayes.

Chúc các bạn ôn thi hiệu quả và đạt kết quả cao trong kì thi!