ĐẠI HỌC THĂNG LONG
Khoa Toán - Tin

ĐỀ CƯƠNG

Lý thuyết xác suất

Sinh viên thực hiện:	**Nguyễn Tú Anh**
MSV:	**A29888**

HÀ NỘI - 2019

Mục lục

Phần I
Biến cố và xác suất của biến cố

1 Biến cố và mối quan hện giữa chúng

Quan hệ giữa các biến cố

- **Kéo theo**: $A \subset B$ - A sảy ra thì B cũng sảy ra.
- **Biến cố đối**: $\overline{A} = \Omega \backslash A$.
- **Biến cố xung khắc**: Không đồng thời xảy ra.

Phép toán tập hợp

- **Hợp**: $A \cup B$ - Sảy ra nếu ít nhất có một trong 2 biến A, B xảy ra.
- **Giao**: AB - Sảy ra nếu cả 2 biến cố A và B sảy ra.

2 Xác suất của một biến cố

- Xác suất đồng khả năng của biến cố A : $P(A) = \frac{|A|}{|\Omega|}$
- Nguyên lý xác suất nhỏ
  - Nguyên lý: Nếu một biến cố có xác suất rất nhỏ thì thực tế có thể cho rằng trong một phép thử biến cố đó sẽ không sảy ra.
  - Kí hiệu : α - mức ý nghĩa($\beta = 1 - \alpha$: Độ tin cậy)

3 Các quy tắc tính xác suất

- Quy tắc cộng xác suất
  - A_1, A_2 .. đôi một xung khắc $\rightarrow P(A_1 \cup A_2 \cup ..) = P(A_1) + P(A_2) + ..$
  - Công thức tổng quát : $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 A_2)$

- Quy tắc chuyển sang biến cố đối: $P(A) = 1 - P(\overline{A})$
- Quy tắc **nhân**: Biến cố A, B **độc lập** $\rightarrow P(AB) = P(A) * P(B)$

4 Phép thử lặp - Công thức Becnuli

- Kí hiệu: $P_k(n; p)$
- Ý nghĩa: Là xác suất để trong một dãy n phép thử **độc lập** biến cố A xuất hiên đúng k lần.
- Công thức: $P_k(n; p) = C_n^k p^k q^{n-k}$
- Chú thích: $p = P(A), q = 1 - p$.

5 Xác suất có điều kiện - Quy tắc nhân tổng quát

- Xác suất của B với điều kiện A : $P(B/A)$
- **Quy tắc nhân tổng quát** : $P(AB) = P(A) * P(B/A)$

6 Công thức xác suất đầy đủ và công thức Bayet

Hệ đầy đủ

Các biến cố $B_1, B_2, ..., B_n$ là một hệ đầy đủ nếu:

- Đôi một xung khắc : $B_iB_j = \varnothing$ với $i \neq j$
- $\Omega = B_1 \cup B_2 ... \cup B_n$: Biến cố chắc chắn

Công thức xác suất đầy đủ

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

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A/B_i)$$

Công thức Bayet

Nếu $B_1...B_n$ là một hệ đầy đủ và biến cố A có $P(A) > 0$ thì mới mỗi $k = 1, 2, ..., n$ ta có:

$$P(B_k/A) = \frac{P(B_k)P(A/B_k)}{P(A)} = \frac{P(B_k)P(A/B_k)}{\sum_{i=1}^{n} P(B_i)P(A/B_i)}$$

$P(B_1), ..., P(B_n)$ là **xác suất tiên nghiệm**. Sau khi biến cố A sảy ra, các xác suất của B_i được tính trên thông tin này- $P(B_i/A)$ được gọi là **xác suất hậu nghiệm**.
Vì thế công thức Bayet còn có tên là **công thức xác suất hậu nghiệm**

Phần II

Đại lượng ngẫu nhiên rời rạc

Một ĐLNN được gọi là **rời rạc** nếu nó chỉ nhận một số hữu hạn các giá trị hoặc một số vô hạn đếm được các giá trị.

1 Phân bố xác suất và hàm phân bố

1.1 Phân bố xác suất

Phân bố xác suất là một bảng như sau:

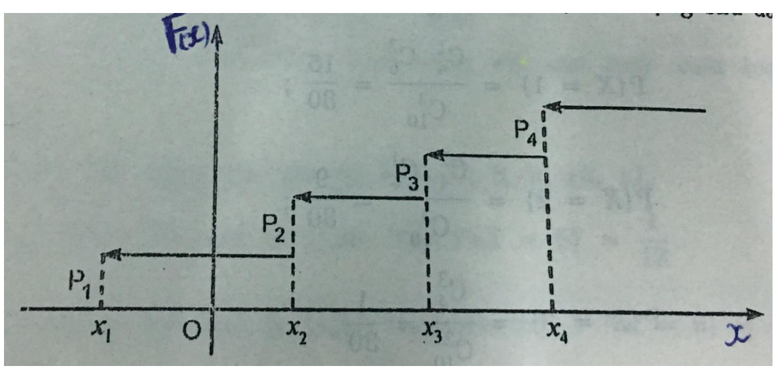
X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Ở đó $p_i = P(X = x_i)$.

1.2 Hàm phân bố

$F(x) = P\{X < x\}$

Hàm phân bố của một ĐLNN rời rạc X là một hàm **bậc thang**, **không giảm** có bước nhảy tại các giá trị có thể của X. Độ lớn của bước nhảy tại điểm x_k là $p_k = P\{X = x_k\}$.



Hình 1: Đồ thị hàm phân bố

1.3 Ví dụ

Bảng phân bố xác suất của X là:

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

Hàm phân bố của X:

$$F(x) = \begin{cases} 0 \text{ nếu } x \leq 0 \\ \frac{5}{30} \text{ nếu } 0 < x \leq 1 \\ \frac{20}{30} \text{ nếu } 1 < x \leq 2 \\ \frac{29}{30} \text{ nếu } 2 < x \leq 3 \\ 1 \text{ nếu } x > 3 \end{cases}$$

2 Kì vọng và phương sai

2.1 Kì vọng

Kì vọng (giá trị trung bình) của X, kí hiệu EX:

$$EX = \sum x_i p_i$$

.

2.2 Phương sai

Phương sai của X, kí hiệu là DX, là độ lệch bình phương trung bình

$$DX = E(X - EX)^2 = EX^2 - (EX)^2 = \sum x_i^2 p_i - (EX)^2 = \sigma^2$$

Ở đó, σ là **độ lệch chuẩn** của X

3 Phân bố đồng thời và hệ số tương quan

3.1 Phân bố đồng thời

Bảng phân bố xác suất đồng thời:

$X \backslash Y$	y_1	y_2 ...	y_j	y_n
x_1	p_{11}	p_{12}	p_{1j}	p_{1n}
x_2	...	...	...	...
	...	...	...	...
x_i	p_{i1}	p_{i2}	p_{ij}	p_{in}
x_m	p_{m1}	p_{m2}	p_{mj}	p_{mn}

Ở đó, $p_{ij} = P(X = x_i, Y = y_j)$

3.2 Đại lượng ngẫu nhiên độc lập

2 ĐLNN X,Y độc lập:

$$P\{X = x_i/Y = y_j\} = P\{X = x_i\}$$

Kí hiệu: $$P_{io} = P\{X = x_i\} = \sum_{j=1}^{n} p_{ij}$$

$$P_{oj} = P\{Y = y_j\} = \sum_{i=1}^{m} p_{ij}$$

X và Y độc lập nếu: $$p_{ij} = p_{io}p_{oj}$$

3.3 Covarian và hệ số tương quan

Covarian

$$cov(X, Y) = \sum_i \sum_j x_i y_j p_{ij} - EX * EY$$

Hệ số tương quan

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

Chú ý: $-1 \leq \rho(X, Y) \leq 1$.

4 Phân bố nhị thức

Quan tâm đến X - số lần xuất hiện biến cố trong n lần thực hiện phép thử.

Kí hiệu: $X \sim B(n, p)$

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

- $EX = np$
- $DX = np(1-p)$

Trong đó: p là xác suất xuất hiện biến cố.

5 Phân bố Poát Xông

Số lần xảy ra thành công của một sự kiện trong 1 khoảng thời gian nhất định.
Kí hiệu: $X \sim P(\lambda)$
Phân bố: $P\{X = k\} = \frac{e^{-\lambda}\lambda^k}{k!}$

- $EX = \lambda$
- $DX = \lambda$

Nếu $X \sim P(\mu)$ và $Y \sim P(\lambda)$ X,Y độc lập thì $(X + Y) \sim P(\mu + \lambda)$

Phần III
Đại lượng ngẫu nhiên liên tục

Một ĐLNN X được gọi là ĐLNN liên tục nếu:
- Tập cáo giá trị có thể của X lấp đầy một hay một khoảng trục số.
- Với mọi số a, $P(X = a) = 0$.

1 Hàm mật độ và hàm phân bố

1.1 Hàm mật độ

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

- $f(x) \geq 0$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(a < X < b) = P(a \leq X \leq b) = \int_a^b f(x)dx$

1.2 Hàm phân bố

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

$$F(x) = P\{X < x\}$$

Tính chất

- $0 \leq F(x) \leq 1$
- Hàm không giảm: Nếu $x_1 < x_2$ thì $F(x_1) \leq F(x_2)$
- $F(x)$ là một hàm liên tục.
- $\lim_{x \to +\infty} F(x) = 1$; $lim_{x \to -\infty} F(x) = 0$
- $f(x) = F'(x)$; $F(x) = \int_{-\infty}^{x} f(t)dt$

2 Kì vọng, phương sai

2.1 Kì vọng

$$EX = \int_{-\infty}^{\infty} xf(x)dx$$

2.2 Phương sai

$$DX = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$$

Với $\mu = EX$

3 Phân bố chuẩn

3.1 Phân bố chẩn tắc

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

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

Hàm phân bố của Z, kí hiệu $\Phi(x)$, là:

$$\Phi(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

Ta có:

- $P\{Z < a\} = \Phi(a)$
- $P\{Z > a\} = 1 - \Phi(a)$
- $P\{a < Z < b\} = \Phi(b) - \Phi(a)$

Kì vọng và phương sai

$EZ = 0$; $DZ = 1$

3.2 Phân bố chuẩn

ĐLNN X được gọi là có phân bố chuẩn với tham số μ và σ^2 nếu ĐLNN $Z = \frac{X-\mu}{\sigma}$ có phân bố chuẩn tắc.
Kí hiệu: $X \sim N(\mu, \sigma^2)$
Ta có:

$$P\{X < x\} = \Phi(\frac{x-\mu}{\sigma})$$
$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$
$$EX = \mu \text{ và } DX = \sigma^2$$

Ta có:

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

Quy luật 68-95-99.7

- $P\{|X-\mu| \leq 1\sigma\} = 68\%$
- $P\{|X-\mu| \leq 2\sigma\} = 95\%$
- $P\{|X-\mu| \leq 3\sigma\} = 99.7\%$
- $P\{|X-\mu| \leq \alpha\sigma\} > 99.99\%$ khi $\alpha > 3$

Tính chất

- $X \sim N(\mu, \sigma^2)$ và a,b là số thực thì $(aX+b) \sim N(a\mu+b, (a\sigma)^2)$
- $X \sim N(\mu_X, \sigma_X^2)$ và $Y \sim N(\mu_Y, \sigma_Y^2)$ với X,Y là các ĐLNN **độc lập** thì:
  - $U = (X+Y) \sim N(\mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2)$
  - $V = (X-Y) \sim N(\mu_X-\mu_Y, \sigma_X^2+\sigma_Y^2)$
  - U và V độc lập

4 Phân phối mũ

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

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

Hàm phân bố là:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} \text{ nếu } x > 0 \\ 0 \text{ nếu } x \leq 0 \end{cases}$$

Kì vọng và phương sai:

$$EX = \frac{1}{\lambda}$$

$$DX = \frac{1}{\lambda^2}$$

5 Phân bố đều

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

Hàm mật độ:

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

Hàm phân bố:

$$F(x) = \begin{cases} \frac{x-a}{b-a} \text{ nếu } x \in [a, b] \\ 0 \text{ nếu } x < a \\ 1 \text{ nếu } x > b \end{cases}$$

Xác suất để X rơi vào (α, β) là:

$$P\{\alpha < X < \beta\} = \frac{\beta - \alpha}{b - a}$$

Kì vọng và phương sai: $EX = \frac{a+b}{2}$; $DX = \frac{(b-a)^2}{12}$

Phần IV

Tính chất quan trọng của tham số đặc trưng

1 Kì vọng

- $E(aX + bY) = aEX + bEY$ với X,Y là 2 ĐLNN bất kỳ và a,b là 2 số thực bất kỳ.
- Nếu X, Y độc lập:$E(XY) = EXEY$.

2 Phương sai

- $D(aX + b) = a^2DX$ với a,b là 2 số thực.
- $DX = E(X^2) - (EX)^2$
- $D(aX + bY) = a^2DX + b^2DY + 2ab * cov(X, Y)$ (X,Y độc lập thì $cov(X, Y) = 0$)

Phần V
Hàm đặc trưng xác suất và các định lý giới hạn

1 Hàm đặc trưng xác suất

Định nghĩa

Cho ĐLNN X, $g(t)$ là hàm đặc trưng của X nếu:

$$g_X(t) = E(e^{itX})$$

Tính chất

- $|g_X(t)| \leq 1 = g_X(0)$
- $g_{aX+b}(t) = e^{itb} g_X(at)$ với a,b là hằng số.
- Dãy $X_1, ..., X_n$ độc lập, $Y = \sum_{k=1}^{n} X_k$ thì $g_Y(t) = \prod_{k=1}^{n} g_{X_k}(t)$

Hệ quả

- $X \sim B(n,p) \Rightarrow g_X(t) = (pe^{it} + (1-p))^n$
- $X \sim P(\lambda) \Rightarrow g_X(t) = e^{\lambda(e^{it}-1)}$
- $X \sim N(0,1) \Rightarrow g_X(t) = e^{-t^2/2}$
- $X \sim N(\mu, \sigma^2) \Rightarrow g_X(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$

Phần VI
Luật số lớn và các định lý giới hạn

1 Các dang hội tụ của dãy các đại lượng ngẫu nhiên

1.1 Hội tụ theo xác suất

Dãy $Z_1, Z_2, ...$ các ĐLNN hội tụ theo xác suất tới ĐLNN Z khi $n \to \infty$ nếu:

$$\forall \varepsilon > 0, P\{|Z_n - Z| > \varepsilon\} \to 0 \text{ khi } n \to \infty$$

1.2 Hội tụ theo bình phương trung bình

Dãy $Z_1, Z_2, ...$ các ĐLNN hội tụ theo bình phương trung bình tới ĐLNN Z nếu:

$$E|Z_n - Z|^2 \to 0 \text{ khi } n \to \infty$$

1.3 Hội tụ theo phân bố

Dãy $Z_1, Z_2, ...$ các ĐLNN hội tụ theo phân bố tới ĐLNN Z nếu:

ĐLNN rời rạc

$$\lim_{n \to \infty} P\{Z_n = c\} = P\{Z = c\}$$

ĐLNN liên tục

$$\lim_{n \to \infty} P\{Z_n < x\} = P\{Z < x\}$$

2 Bất đẳng thức Trê-bư-sép và Luật số lớn

2.1 Bất đẳng thức Trê-bư-sép

Cho Y là ĐLNN không âm, khi đó với mọi $a > 0$ ta có:

$$P\{Y > a\} \leq \frac{EY}{a}$$

Hệ quả: Giả xử X là ĐLNN với $\mu = EX$. Khi đó với mọi $\varepsilon > 0$:

$$P\{|X - \mu| > \varepsilon\} \leq \frac{DX}{\varepsilon^2}$$

2.2 Luật số lớn

Giả xử $X_1, X_2, ..., X_n, ...$ là dãy các ĐLNN **độc lập có cùng phân bố** và có kì vọng là μ và phương sai σ^2. Khi đó $\overline{X} = (X_1 + X_2 + ..X_n)/n$ sẽ hội tụ tới μ theo xác suất.

$$\forall \varepsilon, \lim_{n \to \infty} P\{|\overline{X} - \mu| > \varepsilon\} = 0$$

Định lý Markov: Nếu dãy $X_1, X_2, ...$ thỏa mãn điều kiện:

$$\frac{1}{n^2} D\left(\sum_{i=1}^{n} X_i\right) \to 0$$

khi $n \to \infty$ thì dãy (X_n) tuân theo luật số lớn.

Tài liệu

[1] Đặng Hùng Thắng *Mở đầu về lý thuyết Xác suất và các ứng dụng*, (2011) NXB Giáo dục Việt Nam.