XÁC SUẤT THỐNG KÊ

Phòng 308.

giảng viên : Lê Hoàng Tuấn tuanlh@uit.edu.vn

tài liệu : giáo trình XSTK.

TG Anh Dũng - Huỳnh Minh Trí

Phần mềm hỗ trợ : SPSS , SPLUS

CHƯƠNG MỞ ĐẦU : GIẢI TÍCH TỔ HỢP

1. Một số khái niệm cơ bản :

cho tập hợp $X = \{x_1, x_2, x_n\}$ từ X ta lập ra các nhóm nhỏ hơn gồm có k phần tử → ta gọi là lập mẫu cỡ k

. Vấn đề ta cần quan tâm là tính các số các mẫu có thể được tạo

. giả sử có m công việc

+ CV1 được giải quyết bằng n_1 cách

+ CV2 n_2 cách

+ CV3 n_3 cách

CVm n_m cách

→ ta có $n_1 \times n_2 \times \dots \times n_m$ cách để giải quyết công việc : ta gọi đây là q tắc nhân

Vd1 Từ tập hợp $\{0, 1, 2, 3, 4, 5, 6\}$ lập ra được bao nhiêu số tự nhiên gồm 3 chữ số nếu

a. các chữ số có thể lặp < có lặp / ko có lặp

b. các chữ số không lặp

c. các số lẻ ≥ 303 và ≤ 505

d. các số nguyên tố

e. các số chính phương

f. các số là bội của 3

g. 6

h. các số là số ngũ và ≥ 401 và ≤ 603

i. các số lẻ, có tổng các chữ số là bội của 2 chẵn

j. các số là ước số của 10

k. các số chẵn ≤ 302 hay ≥ 504

e. các số có tổng các chữ số là ước của 720

giải.

cách 1. liệt kê rồi đếm các phần tạo thành

c. gọi số cần tìm là $\overline{abc}$

a. có 3 cách chọn {3, 4, 5}

b. có 7 cách chọn {0, 1, 2, 3, 4, 5, 6}

c. Nếu a là 3 : có 2 cách

câu c. chọn a : 2, b : 7, c : 3, = 42 số − số 301

+ Nếu a = 3. 1 cách + số 501 + 503 + 505

b 7 cách } 21 số = 44 số

c 3 cách } − số 301 = 20 số

+ Nếu a = 4 : 1 cách

b = 7 cách } 21 số

c 3 cách

+ Nếu a = 5 : 1 cách

b 1 cách } 3 số

c 3 cách

i. có 3 phương án cho c.

b thuộc tập $\{0, 2, 4, 6\}$: 4 cách

a $\{1, 3, 5\}$: 3 cách

TH2 : b thuộc tập $\{1, 3, 5\}$: 3 cách

a thuộc tập $\{2, 4, 6\}$ 3 cách

ta có $4.3.3 + 3.3.3 = 63$ số

j. các chữ số chia hết cho 3.

$\Leftrightarrow a + b + c = 3k$.

cho k chạy :

TH1 : $k = 0$.

$k = 1$ ta có $a + k + c = 3$

102, 201, 111, 300, 120, 210 : 6 số

$k = 2$, ta có $a + b + c = 6$.

* Mẫu không lặp

là mẫu mà các phần tử của mẫu chỉ xuất hiện 1 lần

Vd1. Rút ngẫu nhiên 3 lá bài từ bộ bài 52 lá

ta có mẫu cỡ 3 không lặp

Vd2 : Từ 1 lớp 50 SV, chọn ngẫu nhiên 5 người

→ đây là mẫu cỡ 5 không lặp

* Mẫu có lặp : là mẫu mà các phtử của mẫu có thể lặp lại nhiều lần

Vd1. số đt là 1 mẫu có lặp

Vd2. MSSV. 12520212

* **Mẫu không có thứ tự**:
là mẫu mà khi thay đổi vị trí các phần tử khác nhau của mẫu, ta không nhận mẫu mới

vd1: Đó chơi lắc bầu cua: cục xí ngầu là 1 mẫu có 3.

* **Mẫu có thứ tự**: là mẫu mà thay đổi các vị trí của phần tử khác nhau, ta nhận được mẫu mới

Ta có bảng sau:

Mẫu	không lặp	có lặp
có thứ tự	chỉnh hợp	chỉnh hợp lặp
không thứ tự	tổ hợp	tổ hợp lặp

2. Chỉnh hợp chập k từ n phần tử

là mẫu có thứ tự
không lặp

Số chỉnh hợp A_n^k

vd1: giải bóng đá ngoại hạng Anh gồm 20 đội bóng thi đấu với nhau vòng tròn 2 lượt. Hỏi tất cả có bao nhiêu trận

vd2 Từ 1 lớp học chọn ra 3 sv trong 120 sv làm ban cán sự (1LT - 1LP1 (học tập) - 1 LPVN
Hỏi có bao nhiêu cách chọn vào ban cán sự (có thứ tự).

$A_{10}^{3} = 10.9.8$.

Thứ..... Ngày..... Tháng..... Năm.....

a. Nếu không ai nể chết

b. Anh A không làm lớp trưởng $\begin{cases} A \ LP1 \\ A \ LP2 \end{cases}$

thì giữa kì 22/04/2015. $A_{120}^{3} - A_{119}^{1} \xrightarrow{A \ k^{o}} = A_{119}^{1} \cdot A_{119}^{2}$

$A_{n}^{n} = n!$

$A_{n}^{0} = 1$, $0! = 1$.

c. Anh A chỉ chịu tham gia khi chị C tham gia

d. Anh A và chị B chỉ đồng ý cùng tham gia

e. Nếu có chị D tham gia thì phải có anh A

hay: đồng thời có cả 2 cùng đc

hoặc anh E

hoặc: loại trừ, k° có 2 cùng k° đc

f. anh A chỉ làm LP2, nếu chị B làm LP1 ✓

g. có đúng 2 trong 3 anh A, E, F tham gia

nếu có chị C tham gia

h. có ít nhất 1 trong 3 chị B, C, D không

tham gia nếu có anh E tham gia

i. Nếu có anh A thì không có anh E

j. Nếu chị C làm LP1 thì anh A làm LP2

k. Nếu có chị C hay chị D tham gia thì phải

có mặt anh A hoặc anh E

$A_{120}^{3} - 1.2.A_{118}^{1}$ (2: có 2 cách chọn cho E)

l. Nếu anh A làm LT thì anh E không tham gia

m. Không có 2 trong 4 người A, B, C, E tham

gia nếu có mặt anh K

n. Nếu anh K tham gia thì chị C hay chị D

phải tham gia

o. Không ai trong các chị B, C, D, G, H tham gia

làm lớp phó nếu anh A LT

p. Anh A và chị B không tham gia vào cùng

1 ban cán sự

Anh A tham gia thì chị C không tham gia

hoặc nếu chị C tham gia thì chị D không tham gia

Vd3. Từ nhóm 20 sv, chọn 5 sv xếp hàng (có thứ tự).

cách chọn

a. Anh A chỉ chọn đầu hàng 1. A_{19}^{4} (cách)

3. A_{19}^{4}

b. Anh A không đứng đầu hoặc cuối hàng $A_{20}^{5} - 2.A_{19}^{4}$

c. Anh A, chị B chỉ chọn đứng gần nhau: $8.A_{18}^{3}$

d. Anh A chỉ chọn đứng gần chị C

e. Anh A đứng cách chị B không quá 2 ng

f. Anh A và anh D không đứng gần nhau

g. Anh A đứng cách anh D từ 2 x trở lên

h. Anh A đứng vị trí snt nếu chị C đứng vị trí

chẵn

1 cách

i. Vị trí anh A + vị trí chị B = số chính phương

j. Vị trí của anh A + vị trí chị B ≤ ước của 12

k. Vị trí anh A + vị trí chị B ≥ bội của 2

l. Anh A đứng ở vị trí ≤ vị trí của anh E − 2

3. Chỉnh hợp lặp: chập k từ n:

phần tử, là mẫu

có thứ tự

có lặp

$\tilde{A}_{n}^{k} = n \times n \times n \ldots = n^{k}$

nếu p thì q sai trong TH: p đúng q sai

Thứ.....Ngày.....Tháng.....Năm.....

$n \geq k$ hoặc $n < k$

Vd1 từ sđt mobyone với đầu số 090 được đk lập từ 7 kí số. Hỏi thu được bao nhiêu số? 10^7 $\tilde{A}_{10}^{7} = 10^7$ (số)

Vd2. vì sao ASCII có 256 mã : 2^8 (1 byte - 8 bit)

Vd3 Một mã hàng hoá là 1 mẫu có 6 được chọn từ tập {A, K, H, L, O, 1, 3, 5, 8}

Hỏi có thể lập được bao nhiêu mã hàng hoá

Vd4 Một biển số xe gắn máy là dãy kí tự kí số như sau:

2 kí số - 1 kí tự - 1 kí số □□-□□

5 kí số □□□.□□

Hỏi có thể lập được bao nhiêu bảng số xe gắn máy

Vd5 Một biển số xe ô tô là dãy kí tự và kí số

2 kí số - 1 kí tự □□-□

5 kí số □□□.□□

Hỏi có thể đk tối đa bao nhiêu biển số

4. Hoán vị : một hoán vị là 1 mẫu

có thứ tự

không lặp

gồm đủ mặt n phtử

✓ Vd4 xếp chỗ ngồi cho 6 SV vào 6 chỗ. Hỏi có bao nhiêu cách (có thứ tự).

a. nếu anh A chỉ ngồi đầu bàn , 2.5!

b. Anh A ngồi giữa chị B và chị C

c. Anh A chỉ ngồi vị trí chẵn, nếu chị B ngồi vị trí ≥ 3

d. Anh A ngồi cách chị B đúng 2 vị trí

e. Anh A ngồi vị trí snt nếu chị B ngồi vị trí số chính phương $\{1, 4\}$

f. anh A và chị B không ngồi gần

g. vị trí ngồi của anh A + vị trí ngồi của chị B $\geq$ vị trí ngồi chị C

h. Vị trí ngồi của anh A là bội của vị trí ngồi chị B

i. Vị trí ngồi của anh A là ước số của (vị trí ngồi chị B + 2 + vị trí ngồi chị C)

câu 1 Nếu vận đầu là 1 mẫu số 2 phần 2

+ số khác nhau

+ không lặp

→ Đây là chỉnh hợp không lặp chập 2 của 20

$A_{20}^{2} = 20 \times 19 = 380$. (mẫu).

Thứ.....Ngày.....Tháng.....Năm.....

Vd2. câu c: Anh A chỉ tham gia khi chị C tham gia.

→ Nếu C tham gia thì A tham gia

C tham gia → A tham gia. $2.A^1_{118}$.

ta có tg mẫu lớn: A^3_{120} cách chọn bất kì.

trừ cho trường hợp: C tham gia A k° tham gia

$A^3_{120} - 3.A^2_{118}$.

C tham gia A tham gia. $A^2_3.A^1_{118}$

C tham gia A k° tham gia $A^1_3.A^2_{118}$

C không tham gia → C không tham gia

A^3_{118}

d. TH1 A tham gia, B không tham gia: $3.A^2_{118}$

A k° tham gia, B tham gia: $3.A^2_{118}$

→ $A^2_{120} - 6.A^2_{118}$

e. TH1. A tham gia - E tham gia: ~~$3 - A^2_{117}$~~ 3! cách

A không tham gia - E không tham gia. $3.A^2_{11}$

Lấy không gian mẫu - TH1 - TH2.

Vd1.

k. có 1 trong 2 chị C hay D

cả 2 A và E đều tham gia. P_3.

cả 2 đều không tham gia $A^1_3.A^2_{116}$

+ ~~cả 2 A và E đều tham gia~~

có 2 trong chị C và D tham gia.

+ không có A và E tham gia. $A^2_3.A^1_{116}$

cả 2 A và E tham gia. (vô lí).

- Số cách $A^3_{120} - 2(P_3 + A^1_3.A^2_{116}) - A^2_3.A^1_{116}$

vd4 câu e:

TH1 có 2 cách xếp B.

3 cách xếp A

4! các cách xếp còn lại

$\Big\} \Rightarrow$ có $2.3.4!$

TH2: B không ngồi vị trí chính phương ra có: $4.5!$

5. Tổ hợp chập k từ n phần tử là mẫu không không có thứ tự

+ không lặp $\Big\}$ có k

- số tổ hợp $C_n^k = \dfrac{A_n^k}{k!}$

hay $C_n^k = \dfrac{n!}{k!(n-k)!}$

vd1 Bài hộp bi có 8 đỏ, 12 xanh, 6 vàng, 14 đen lấy 1 lần 7 bi. Bình số trường hợp.

a. khải có bi đỏ

b. 2 bi đỏ, ít nhất 2 bi xanh

c. ít nhất 2 bi đỏ, nhiều nhất 2 xanh, tối đa 2 vàng

d. không có bi đỏ nếu có từ 2 bi đen trở lên

e. bi đỏ $\geq$ bi xanh $\geq$ bi vàng $\geq$ bi đen

f. $\leq$ $\leq$ $\leq$

g. có đúng 4 bi cùng màu

h. có ít nhất 5 bi cùng màu

i. có đúng 2 màu bi

j. có ít nhất 3 màu bi

k. Nếu có bi đỏ thì không có bi đen, nhưng phải có bi xanh

l) bi đỏ ≥ (bi xanh + bi vàng)

m) bi đỏ là bội số của bi đen

o) số bi xanh là ước của (bi vàng + bi đỏ).

p. số bi đen là số ngũ nếu bi đỏ là số chính phương

q. số bi xanh là số chẵn nếu bi vàng ≤ bi đen

r. Nếu bi đen là bội số của 3 thì phải có bi đỏ nhưng không có bi xanh

s. Không có bi đỏ nếu có ít hơn 2 bi đen

t. số bi đỏ gấp 2 số bi xanh và số bi vàng nhiều hơn số bi đen

u. Không có ít nhất 2 trong 3 bi (đỏ, vàng, đen) nếu số bi xanh > 5.

v. Phải có ít nhất 2 trong 3 màu bi (vàng, xanh, đỏ) nếu số bi đen ≤ 3.

w. số bi đen ít hơn số bi vàng nhưng nhiều hơn số bi xanh

x. Không có bi đen thì không có bi xanh nếu số bi đỏ ≤ 2.

y. số bi đen = 3 * số bi vàng nếu không có bi xanh

z. số bi đỏ ≤ số bi vàng − số bi xanh + 1 nếu số bi đen là số chẵn

Vd.2 Rút ngẫu nhiên 6 lá từ bộ 52 lá. Tính số TH xảy ra

a. có đúng 2 lá ách - 1 lá già

b. ít nhất 2 lá ách, ít nhất 2 lá già - phải có lá đầm

c. Không có lá ách, không có lá tây

d. có đúng 4 lá rô

e. có ít nhất 4 lá cơ

f. có đúng 2 loại bài

g. có ít nhất 3 loại bài

h. có đúng 4 lá cùng loại

i. có ít nhất 4 lá cùng loại

j. nếu có lá cơ thì k° có lá chuồn

k. số lá cơ ≥ số lá chuồn ≥ số lá bích

l. có ít nhất 2 lá cơ, nhiều nhất 2 lá rô và tối đa 2 lá bích

m. lá cơ có số chẵn nếu số lá bích ≤ số lá chuồn

n. có ít nhất 2 loại bài (cơ, rô, chuồn) nếu số lá bích ≤ 2

o. Không có lá bích nếu số lá cơ gấp 2 lần số lá chuồn

p. số lá cơ + lá chuồn ≤ số lá rô

q. số lá cơ = 3 số lá bích - 2 số lá chuồn

r. có đúng 5 lá liên tiếp cùng loại và không có lá tây

s. có ít nhất 5 lá cùng loại và k° có lá tây

t. có ít nhất 5 lá đỏ

u. có nhiều nhất 4 lá đen

v. số lá đỏ ≥ số lá đen

w. số lá cờ + số lá rô = số nghé

x. số lá cờ là ước của 12 nếu số lá rô là bội của 5

Vd3. Từ 1 nhóm 48 sv nam và 22 sv nữ, ta chọn ra 1 đội 8 người. Tính số TH xảy ra

a. Anh A không tham gia

b. Anh A không tham gia nếu anh B tham gia

c. A chỉ tham gia nếu C hay chị D tham gia

d. A chỉ tham gia nếu nữ ≥ 4

e. A tham gia khi có đúng 2 người là C và D tham gia và không có B

f. số nữ là số chính phương

g. số nữ ≥ số nam

h. phải có anh A, chị C nếu số nữ ≥ số nam

i. Phải có anh A nếu số nữ là bội số của số nam

p. Nếu có chị C thì chị D không tham gia

k. Nếu có mặt anh A thì không có nữ nào tham gia

l. Phải có anh A và anh B, đồng thời không có 2 trong 3 chị (C, D, E) tham gia.

m. Phải có chị C và chị D nếu số nữ là snt

n. Không có anh A hoặc không có anh B nếu có ít hơn 3 nữ

o. có ít nhất 3 nam và tối thiểu là 2 nữ nếu có anh A và chị B tham gia

p. Phải có anh A nếu có chị C và chị D, nhưng không có anh B và không có anh F

q. Không có ít nhất 2 trong 4 anh (A, B, F, G) nếu không có chị C tham gia

r. có ít nhất 2 chị trong số (C, D, F, H) không tham gia nếu có anh F hoặc anh B

s. Phải có anh A nếu số nữ ≥ 3 lần số nam

t. Phải có chị C nếu có anh A và có anh B và số nữ là số ngũ

u. Phải có chị C nếu số nữ là bội của (số nam + 2)

v. Nếu có anh A thì phải có chị C và chị D nhưng không có chị E và số nam ≥ số nữ

w. Không có hai trong số các chị C, D, E, H tham gia nếu số nam ≤ 3.

x. Có đúng 3 anh trong số A, B, F, G không tham gia nếu k° có ít 3 nữ trở lên

y. Phải có anh A và có chị C nếu có ít nhất 2 trong 4 anh B, F, G, K tham gia, và nhiều nhất 4 nữ

z. Không có đúng 2 anh trong số A, B, F, K tham gia

6. Tổ hợp lặp: chập k từ n phtử là mẫu

+ không thứ tự

+ có lặp

số tổ hợp lặp $\tilde{C}_n^k = C_{n+k-1}^k$

$n \geq k$

Vd4 Phát 5 bạc lưỡng giống nhau cho 24 sv. Hỏi có bao nhiêu cách phát (nếu k° có ai từ chối) mẫu có 5, k° thứ tự ~ để hoán k° càng hoá. mẫu có lặp (do 1 sv nhận n' tb) C_{24}^{5}

Vd5 xếp 42 quyển sách vào 1 kệ sách có 4 ngăn. Tính số trường hợp xảy ra, với (tổng hợp) $= C_{24+5-1}^{5}$

(Vd5: mẫu có 4l: chia 42 sách vào 4 ngăn)

a. Ngăn 1 ít có sách: ngăn 1 k° sách $= C_{28}^{5}$

b. Ngăn kệ nào cũng có sách

c. có ít nhất 2 ngăn k° có sách

d. Ngăn nào cũng có từ 9 quyển trở lên

e. số sách ngăn 1 ≥ số sách ngăn 2 ≥ ngăn 3 ≥ ngăn 4

f. Ngăn một không có sách nếu ngăn 2 có từ 10 quyển

g. số sách ngăn 2 gấp 3 số sách ngăn 4

h. số sách ngăn 1 ít hơn số sách ngăn 2, nhiều hơn số sách ngăn 3

i. số sách ngăn 1 là bội số (số sách ngăn 3 + số sách ngăn 4)

j. số sách ngăn 1 là snt

k. số sách ngăn 3 là số chính phương nếu số sách ngăn 2 ít hơn số sách ngăn 1

* Phép lấy mẫu có lặp. để lấy mẫu có lặp ta làm như sau: chọn phử thứ nhất, ghi nhận rồi trả lại rồi lặp lượt sau đó cứ tiếp tục như vậy cho đến khi đủ cỡ.

* Phép lấy mẫu k° lặp: chọn phử thứ 1, ghi nhận rồi bỏ

ra ngẫu nhiên

CHƯƠNG I: SỰ KIỆN VÀ XÁC SUẤT

1. Đối tượng nghiên cứu và lí thuyết xác suất

Theo quan điểm của toán học thống kê thì trong tự nhiên có 2 loại hiện tượng.

Hiện tượng tất nhiên

Hiện tượng ngẫu nhiên

Đối tượng nghiên cứu là hiện tượng ngẫu nhiên

Mục đích: giúp ta tìm ra qui luật biến động của hiện tượng bằng cách lập mô hình toán học cho các hiện tượng đó rồi đó xđinh cho giá trị thông số thu được

2. Sự kiện và phép thử:

+ phép thử: có thể xảy 1 thí nghiệm nào đó, có thể quan sát 1 hiện tượng nào đó

+ sự kiện: là kết quả của phép thử và ta thường kí hiệu A, B, C, D

Vd1 tung 1 đồng xu là 1 phép thử với sự kiện KQ là

S = { đồng xu xuất hiện mặt sấp }

N = { ngửa }

Vd2 tung một con xúc xắc thì ta có các sự kiện là

1 = { xuất hiện mặt 1 chấm }

A = { xuất hiện chấm lẻ }

B = { xuất hiện mặt có số chấm không là số chính phương }

(Ω còn được gọi là kg mẫu).

* Sự kiện chắc chắn: là sự kiện nhất định xảy ra khi ta thực hiện phép thử và ta kí hiệu là Ω

Vd. Tung 1 đồng xu thì Ω = {S, N}

tung 1 con xúc xắc thì Ω = {1, 2, 3, 4, 5, 6}

* Sự kiện không thể: là sk nhất định không xảy ra khi ta thực hiện phép thử và kí hiệu là ϕ (phi)

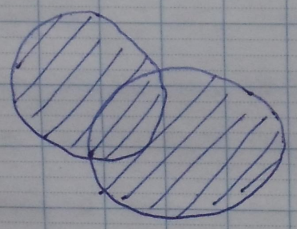
Vd. Tung 1 đồng xu 5 cạnh thì ϕ = {

3. Mối quan hệ giữa các sk:

ⓐ. Sự kiện tổng. (hiểu là hết).

A + B hay A ∪ B đgl tổng của hai sự kiện A và B.

Là sự kiện xảy ra ⟺ ít nhất 1 trong chúng xảy ra



A + B.

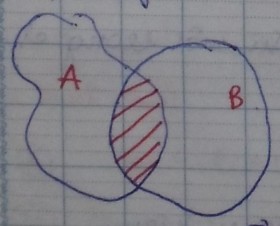
Vd. A = {2, 3, 5, 4, 0}

B = {1, 3, 5, 7}.

⇒ A + B = {

ⓑ Sự kiện tích (giao là giống).

AB hay A ∩ B đgl tích của 2 sự kiện A và B. Là sự kiện xảy ra ⟺ cả A và B cùng xảy ra



A ∩ B.

A = {1, 4, 7, 2, 5}

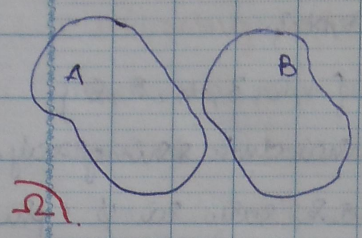
B = {3, 1, 4, 8, 0, 9}

⇒ AB =

ⓒ. Sự kiện xung khắc: không có cái gì chung.

A và B đgl 2 sự kiện xung khắc nhau nếu chúng không thể cùng xảy ra nếu $AB = \phi$

$(A + B \neq \Omega)$.

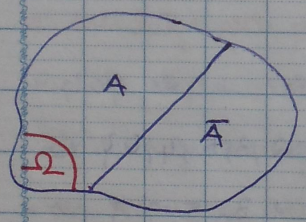


d. Sự kiện đối lập : $\bar{A}$ đgl sự kiện đối lập của A là sự kiện xảy ra $\Leftrightarrow$ A không xảy ra

ta kí hiệu là $\bar{A}$

$$\begin{cases} A.\bar{A} = \phi \\ A + \bar{A} = \Omega \end{cases}$$

(Khi thực hiện phép thử thì 1 trong chúng phải xảy ra)



e. Nhóm đầy đủ:

các sự kiện $A_1, A_2, \ldots, A_n$ đgl nhóm đầy đủ nếu

+ chúng xung khắc từng đôi một

$$A_i . A_j = \phi.$$

+ Khi thực hiện một phép thử thì một trong chúng phải xảy ra, nghĩa là

$$A_1 + A_2 + \ldots + A_n = \Omega.$$

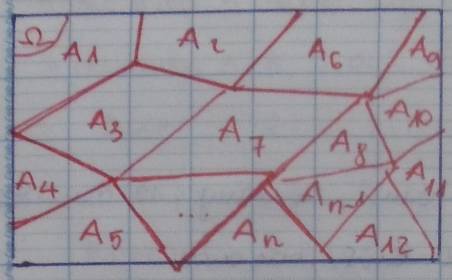
Vd. tung 1 con xúc xắc ra gọi các sự kiện

$A = \{$ xuất hiện mặt có số chấm là số nguyên $\}$.

B = { xuất hiện các mặt ra số chính phương }

C = { xuất hiện các mặt có số chấm ra bội của 6 }

CMR A, B, C tạo thành nhóm đầy đủ



4. Xác suất (probability).

Vd. Từ một hộp có 8 đỏ, 8 xanh, 10 vàng. ta lấy ngẫu nhiên 3 bi (tổ hợp < không lặp / không thứ tự >).

gọi A = { lấy được 3 bi đỏ } C_{12}^3

thì ta thấy khả năng xảy ra của A là C_8^3

B = { lấy được 3 bi xanh } thì khả năng xuất hiện của B là . C

C = { lấy được 3 bi 3 màu ≠ nhau } $C_{12}^1 \times C_8^1 \times C_{10}^1 = 960$.

thì khả năng xảy ra của C là .

D = { lấy ít nhất 2 đỏ } < TH1 : $C_{12}^2 \times C_{18}^1$ / TH2 : 3 bi đỏ : C_{12}^3) 1408) $C_{12}^2 \times C_{28}^1$ (Sai)

* ta gọi khả năng xảy ra của sự kiện A là ~~xác~~ suất của A, kí hiệu P(A). (sai $C_{12}^2 \times C_{28}^1$).

* Tính chất. (lấy 4 con số sau dấu phẩy).

(1) $0 \le P(A) \le 1$.

(2) $P(\bar{A}) = 1 - P(A)$

(3) $P(\Omega) = 1$

(4) $P(\phi) = 0$.

(5) $P(A+B) = P(A) + P(B)$ nếu A, B là hai biến cố xung khắc

* Kí hiệu :

[A] tập hợp các trường hợp có thể xảy ra với A

[Ω] : không gian mẫu

thì ta có :

$$P(A) = \frac{[A]}{[\Omega]}$$

Vd. từ 1 hộp 18 bi đỏ, 6 xanh, 14 vàng, 12 đen. Lấy ngẫu nhiên 1 lần 4 bi. Tính xác suất

a. có ít nhất 2 đỏ - nhiều nhất 1 bi xanh

b. có ít nhất 2 màu bi

c. Màu đỏ và màu xanh không cùng xuất hiện

d. số bi đỏ > số bi xanh nếu có bi vàng

e. tối đa 2 đỏ nhiều nhất 1 đen nếu không có bi vàng

5. Định nghĩa xác suất theo qđiểm hình học

giả sử A có thể biểu diễn bằng miền hình học D_A có dtích là S_A

giả sử Ω có thể biểu diễn bằng miền D_Ω có dtích S_Ω

thì xác suất của A là $\frac{S(A)}{S D(\Omega)}$.

6. công thức tính xác suất

a. ct cộng xác suất

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

Nếu $AB = \phi \rightarrow P(AB) = \phi$

→ $P(A+B) = P(A) + P(B)$.

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

b. sự độc lập

2 sự kiện A và B dgl độc lập nếu $P(AB) = P(A).P(B)$

. 3 sự kiện A, B, C dgl độc lập với nhau

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

$P(AC) = P(A).P(C)$.

$P(BC) = P(B).P(C)$

⇒ $P(ABC) = P(A).P(B).P(C)$

c. công thức xác suất có điều kiện

- xác suất của sk A với đk B xảy ra, kí hiệu

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

d. công thức nhân xác suất : xảy ra theo thứ tự

$P(AB) = P(A).P(B|A)$ nếu A xảy ra trước

$= P(B).P(A|B)$ nếu B xảy ra trước

. $P(ABC) = P(A).P(B|A).P(C|AB)$.

$= P(A).P(C|A).P(B|AC)$.

$= P(B).P(A|B).P(C|BA)$.

$= P(B).P(C|B).P(A|BC)$

$= P(C).P(A|C).P(B|CA)$

$= P(C).P(B|C).P(A|CA)$.

e. công thức xác suất toàn phần

Cho $A_1, A_2, \dots, A_n$ là 1 nhóm đầy đủ và A là sự kiện cần qtâm

ta có:

$$P(A) = P(A_1).P(A|A_1)$$
$$P(A_2).P(A|A_2).$$
$$+ \dots + P(A_n).P(A|A_n).$$

Ngoài ra, ta còn có công thức Bayes:

$$P(A_j|A) = \frac{P(A_j.A)}{P(A)} = \frac{P(A_j).P(A|A_j)}{P(A)}.$$

* Sửa bài tập.

B8/trang 6.

gọi A là biến cố: người đó chết

→ $\bar{A}$ là biến cố: người đó không chết

ta có A và $\bar{A}$ lập thành 1 nhóm đầy đủ

a. ta có xác suất để 1 đó không chết

$$P(\bar{A}) = 1 - P(\bar{A}) = 0,9$$

b. gọi G là biến cố: người đó bị tật

ta có $P(G) = P(A+B) = P(A) + P(B) = 0,5)$

$$P(G|\bar{A}) = \frac{P(G).P(\bar{A}|G)}{P(\bar{A})} = \frac{0,5}{0,9} = \frac{5}{9}$$

B26. gọi A là biến cố : { Đội VN }.

B là biến cố { Đạo Nữ }.

C là biến cố { Đạo Pháp }

Ba biến cố đầy đủ.

Câu 7. gọi A là biến cố : xạ thủ A bắn trúng

B là biến cố : xạ thủ B bắn trúng

C là biến cố : xạ thủ C bắn trúng

b. có hai viên đạn trúng địch

$D = \bar{A}BC + A\bar{B}C + AB\bar{C}$.

Biến cố độc lập - xung khắc.

c. gọi E là biến cố có ít nhất 1 viên đạn trúng

$\Rightarrow$ $\bar{E}$ là biến cố không có viên đạn nào trúng

$\Rightarrow$ $\bar{E} = \overline{ABC}$

d. Duy nhất 1 viên

$F = \bar{A}.\bar{B}.C + A.\bar{B}.\bar{C} + \bar{A}.B\bar{C}$

Câu 7. gọi A_i là biến cố : người thứ i bắn trúng

d. gọi D là biến cố có 2 viên trúng địch

$P(\bar{A}|A)$.

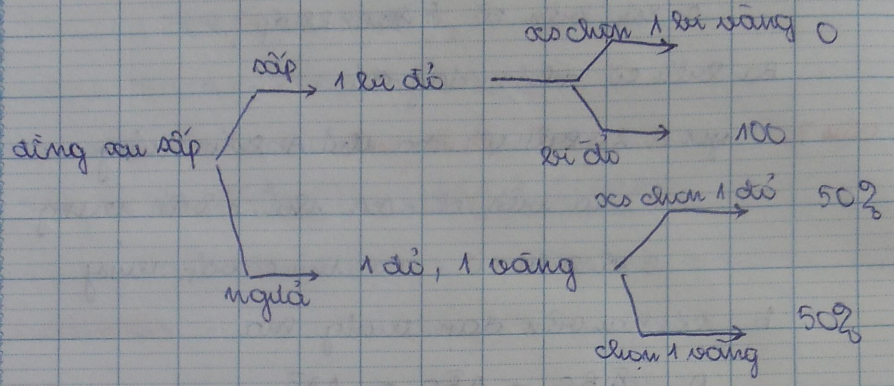
$P(\bar{A}|D)$ xác suất của A bắn trật biết có 2 viên trúng mục tiêu (A trật và có 2 viên trúng).

Câu 9.

gọi A_i là biến cố kiểm tra ở lần thứ i phát hiện 3 lỗ sai lệch

câu 12. A = { sấp }

B = { ngửa }



gọi C là biến cố rút đc 1 vàng.

$P(C) = P(A).P(C|A) + P(B).P(C|B)$.

Bài 29. gọi A_i là biến cố sinh viên ở tổ thứ i

a. gọi A là biến cố chọn đc

Btập Một thùng sản phẩm gồm 14 bóng đèn màu đỏ 18 bóng xanh + 16 bóng vàng + 12 đèn tím. Có 1 khách hàng đến mua sp bằng cách lấy lần lượt từ 3 bóng thì dừng lại. Người này chỉ đồng ý mua hàng nếu trong 3 bóng lấy ra có đúng 2 màu

a. tính xác suất để khách này mua đc hàng

b. giả sử người thứ hai đến mua hàng. Người này cũng lấy lần lượt 3 bóng đèn. Nếu có số đỏ không nhiều hơn số bóng xanh thì ý này đồng ý mua hàng. tính xác suất ý này mua đc hàng

1 bán quên rằng đã bán bóng màu gì cho khách trước

CHƯƠNG II: BIẾN NGẪU NHIÊN

1. Một số khái niệm cơ bản

BNN X là một ánh xạ $X: \Omega \to \mathbb{R}$

hay nói cách khác. biến ngẫu nhiên là loại biến nhận giá trị của mình 1 cách ngẫu nhiên, phụ thuộc kết quả khi ju thuộc phép thử

vd: tung hai đồng xu cùng lúc nếu có 1 mặt sấp thì ta thắng 4000 ngược lại nếu có 1 mặt ngửa thì ta thua 3000

gọi X là số tiền thu được → X là BNN

BNN rời rạc có tập hợp giá trị là 1 tập hợp: rời rạc, hữu hạn và đếm đc

BNN liên tục: có tập hợp các giá trị là 1 tập hợp: liên tục, không đếm được, có vô hạn phần tử, thường được biểu diễn là 1 đoạn thẳng trên trục số

2. Bảng phân phối xác suất của biến ngẫu nhiên rời rạc:

là bảng có dạng

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

Trong đó x_i là các giá trị của BNN X

$p_i = P(X = x_i)$ = xs để X nhận giá trị X nhận giá trị x_i

ta có $p_0 + p_1 + \dots + p_n = 1$.

3. Hàm phân phối xác suất (Function distribution)

Hàm pp xác suất của X là

$F(x) = P(X \leq x) \quad \forall x \in R$

$P(-\infty < X \leq x)$

là xác suất để X nhận giá trị trên khoảng $(-\infty, x]$ khi ta thực hiện phép thử

* tính chất.

$0 \leq F(x) \leq 1$ do $F(x)$ cũng là xác suất

(2) $F(x)$ là hàm không giảm, nghĩa là

$F(a) \leq F(b), \quad a \leq b$

(3) $P(a \leq X \leq b) = F(b) - F(a)$

(4) $\lim_{x \to \infty} F(x) = F(+\infty) = 1.$

$\lim_{x \to -\infty} F(x) = F(-\infty) = 0.$

4. Hàm phân phối xác suất của BNN liên tục

Hàm mật độ của BNN liên tục X là:

$f(x) = F'(x), \quad \forall x \in R$

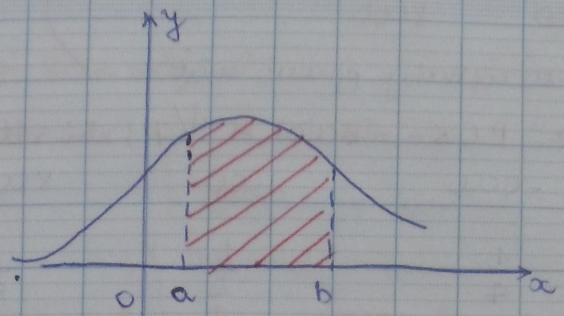
* tính chất.

(1) $f(x) \geq 0$ do $f(x)$ là đạo hàm của các hiến không giảm

(2) $F(x) = \int_{-\infty}^{x} f(t)\,dt$ (tích phân suy rộng).

x là 1 cận chạy

3) $P(a \leq X \leq b) = \int_a^b f(x)\,dx$



☆ (4) $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Vd1. Hãy lập bảng phân phối xác suất cho X (ở ví dụ ở phần 1, sau đó tìm f(x), vẽ đồ thị f(x)

Vd2. Một vận động viên bóng rổ ném 5 quả vào rổ, cho đến khi trúng rổ hoặc hết bóng rổ thì dừng lại. gọi X là số lần đã ném. Biết rằng khả năng ném trúng rổ của VĐV này là 80%

muốn X = 1 thì lần này phải trúng.

X = 2 lần 1 trật lần 2, trúng

X = 5 trúng cũ dọc hoặc trật cũ dọc

Hãy lập bảng pp xác suất và tìm F(x)

* hướng dẫn vd1

các trường hợp khi tung 2 đồng xu

TH1. SS hai mặt sấp → +8000đ

→ $P(X=8) = P(SS) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

TH2. 1 sấp - 1 ngửa → X = 4000 - 3000 = 1.000

⇒ $P(X = 1000) = P(SN) = \frac{1}{4}$

TH3. 1 ngửa - 1 sấp : → X = 4000 - 1000 = 3000

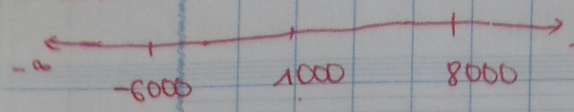
⇒ $P(X = 3000) = P(NS) = \frac{1}{4}$

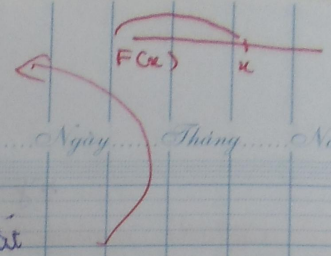
TH4. NN ⇒ X = - 6000

Tiếng học lỗ Châu học văn QT BOOK

tiếp theo, tìm hàm phân phối $F(x)$

$= P(X \le x)\ \forall x \in R \Rightarrow F(x) = 0$





Thứ..... Ngày..... Tháng..... Năm.....

khi cho nên, ta có bảng pp xác suất

$-6000 \le x < 1000 \Rightarrow P(X = -6000) = 1/4$; chỉ xét giá trị < 1000

khi	x	-6000	1000	8000
$1000 \le x < 8000$		$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$P(X = -6000)$ ~~...~~ $P(X \ge 1000) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$

khi $x \ge 8000$ cho x là biến NN liên tục có pp đều trên đoạn từ

$F(x) = P(X) \in [6, 10]$ và ta kí hiệu $x \sim U[6, 10]$. Hãy chỉ ra

$= 1$. hàm mật độ, sau đó chỉ ra hàm pp $F(x)$

lấy tính $P(6 \le X \le 7)$

Như vậy $F(x) = \begin{cases} 0 & \text{khi } x < -6000 \\ \frac{1}{4} & \text{khi } -6000 \le x < 1000 \\ 3/4 & \text{khi } 1000 \le x < 8000 \\ 1 & \text{,} \quad x \ge 8000 \end{cases}$

Bt4. Cho x là một BNN liên tục, có hàm mật độ

$$f(x) = \begin{cases} A(x^2 \ln x + x e^{-x^2}) & x \in [1, 2] \\ 0 & \text{nói } \neq \end{cases}$$

a. tìm A

b. tìm $F(x)$

c. tính $P(1 \le X \le \frac{7}{4})$

Cho x là 1 BNN liên tục có hàm mật độ

$$f(x) = \begin{cases} A e^{-2x} & \text{khi } x \ge 0 \\ 0 & \text{khi } x < 0 \end{cases}$$

qui ước hàm mật độ của phân phối đều.

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

vd5: ta có $\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{+\infty} A e^{-2x} dx$

$= 0 - \frac{1}{2} A . e^{-2x} \Big|_{0}^{+\infty} = \frac{-1}{2} A \cdot (-1) = 1.$

$\rightarrow A = 2$

b. ta có $F(x) = \int_{-\infty}^{x} f(t)dt$

trường hợp 1. $x<0$: ta có $F(x) = \int_{-\infty}^{0} 0\,dt = 0$

(do k° có hàm mật độ → k° có hàm phân phối).

trường hợp 2. $x \geq 0$ ta có

$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} A e^{-2t} dt.$

$= \int_{0}^{x} 2 . e^{-2t} dt = -e^{-2t} \Big|_{0}^{1} = 1 - e^{-2x}$

vậy $F(x) = \begin{cases} 0 & \text{khi } x<0 \\ 1 - e^{-2x} & \text{khi } x \geq 0. \end{cases}$

c (0 ≤ x ≤ 4) $\int_{0}^{4} f(x)dx = \int_{0}^{A} 2e^{-2x} dx$

$= F(4) - F(0)$

ct của hàm pphối

$f(x) = F'(x)$

đạo hàm pphối → mật độ

→ nguyên hàm mđ

* sửa bt tập:

Thứ..... Ngày..... Tháng..... Năm.....

* câu 15:

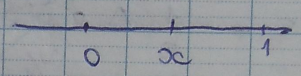
ta có hàm mật độ

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

. $x < 0$

ta có $\int_{-\infty}^{0} 0\,dt = 0$.

. $0 < x < 1$



$$\int_{-\infty}^{0} 0\,dt + \int_{0}^{x} 1\,dx = t\Big|_{0}^{x} = x.$$

. $x > 1$.

$$\int_{-\infty}^{0} 0\,dt + \int_{0}^{1} 1\,dt + \int_{1}^{x} 0\,dt = 1.$$

$$F(x) = \begin{cases} 0 & \text{nếu } x < 0. \\ x & \text{nếu } 0 \le x \le 1 \\ 1 & \text{nếu } x > 1 \end{cases}$$

Chương 2:

* Bài 16.

$G(y) = P(Y \le y) = P(5\ln X \le y)$.

xét $-5\ln X = y$.

$\rightarrow x = e^{-y/5}$

Vậy $\int_0^{e^{-y/5}} dx = e^{-y/5}$

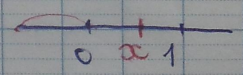
ta có hàm phân phối $G(y) = e^{-y/5}$

Gọi $f(x)$ và $F(x)$ lần lượt là hàm mật độ và hàm phân phối của BNN X

gọi $f(y)$ và $F(y)$ lll hàm mật độ, hàm phân phối của BNN Y

do $X \sim [0,1]$

nên $f(x) = \begin{cases} \frac{1}{1-0} & \text{khi } x \in [0,1] \\ 0 & \text{khi } x \notin [0,1] \end{cases}$



tiếp theo ta tìm hàm phân phối như sau

ta có: $F(x) = F(X \le x)\ \forall x \in \mathbb{R}$

TH1. Khi $x < 0 \Rightarrow \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0\,dt = 0$

Khi $0 < x < 1$

$\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} 1\,dt = t\Big|_0^x = x$

Khi $x > 1$: $\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0\,dt + \int_{0}^{1} 1\,dt + \int_{1}^{x} 0\,dt$

Tiên học lễ, Hậu học văn QT BOOK

$= 1$.

$P(X \le \square) = F(\square)$

$e^{\square} > 0 \ \forall \square$

Thứ.....Ngày.....Tháng.....Năm.....

vậy $F(x) = \begin{cases} 0 & \text{khi } x < 0 \\ x & \text{khi } 0 \le x \le 1 \\ 1 & \text{khi } x > 1 \end{cases}$

ta có $Y = \ln X$

ta có ~~G(y)~~ $= G(Y \le y)$

$= P(-5\ln X \le y)$

$= \cancel{P(X \le e^{-y/5})}$

ta có $G(y) = P(Y \le y)$

$= P(-5\ln X \le y)$

$P(\ln X \ge -\frac{y}{5})$

$P(X \ge e^{-y/5})$

$= 1 - P(X < e^{-y/5})$

$= 1 - F(e^{-y/5})$

ta có: $Y = -5\ln X$

$\Rightarrow$ tương ứng: $y = -5\ln x$

$\Rightarrow \ln x = -\frac{y}{5}$

$\Rightarrow x = e^{-y/5}$

mà x có nghĩa khi $x \in [0,1]$

$\Rightarrow e^{-y/5} \in [0,1]$

mà $F(x) = x$ khi $x \in [0,1]$

$\Rightarrow F(e^{-y/5}) = e^{-y/5}$

vậy

$G(y) = \begin{cases} 1 - e^{-y/5} & \text{khi } e^{-y/5} \in [0,1] \\ 0 & \text{khi } e^{-y/5} \notin [0,1] \end{cases}$

Một trang vở một tương lai QT BOOK

ta có: $0 < e^{-y/5} \leq 1$.

$\Rightarrow e^{-y/5} \leq e^{0}$.

$\Rightarrow -y/5 \leq 0 \Rightarrow y > 0$.

ta có:

$$G(y) = \begin{cases} 1 - e^{-y/5} & \text{khi } y > 0. \\ 0 & \text{khi } y \leq 0. \end{cases}$$

$$\Rightarrow g(y) = G'(y) = \begin{cases} \frac{1}{5} e^{-y/5} & \text{khi } y > 0. \\ 0 & \text{khi } y \leq 0. \end{cases}$$

5. Các đặc trưng của BNN

a. Kì vọng

TH1: X là BNN rời rạc

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

$\Rightarrow E(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n$

TH2: X là BNN liên tục

$$E(X) = \int_{-\infty}^{+\infty} x . f(x)\, dx$$

Expectation

* Ý nghĩa: Kì vọng là giá trị trung bình lí tưởng khi số lần thực hiện phép thử tăng vô hạn

trung thực rồi thì có thể xác kì vọng mà thôi

* Tính chất:

1. $E(c) = c$ với c là hằng số
2. $E(c.X) = c.E(X)$ với c là hằng số
3. $E(X+Y) = E(X) + E(Y)$
4. $E(X.Y) = E(X).E(Y)$

nếu X và Y độc lập với nhau

5. $g(x)$ là 1 hàm số nào đó

TH1 X là BNN rời rạc

X	x_1	x_2		x_n
g(X)	$g(x_1)$	$g(x_2)$		$g(n)$
P	P_1	P_2		P_n

$\Rightarrow E(g(X)) = g(x_1)P_1 + g(x_2)P_2 + \dots + g(x_n)P_n$

TH2: X là BNN liên tục

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

Vd1 Cho X và X là BNN rời rạc có bảng pp sau

X	-4	0	1	3
	0,2	0,4	0,3	0,1

Y	-2	1	4	5
	0,3	0,25	0,15	0,3

a. Tìm E(X) và E(Y)

b. Tìm E(X+Y) và E(X.Y)

c. Tìm $E(4X^2 - 5Y^3)$

d. Tìm $E(Y^4 - 3X^2)$

Vd2. X là BNN liên tục có hàm mật độ

$$f(x) = \begin{cases} A e^{-4x} & \text{khi } x \geq 0 \\ 0 & x < 0 \end{cases}$$

a. Tìm A.

b. Tìm $E(X)$ và $E(X^2)$

c. Tìm $E(4X^2 - 3)$

d. Tìm $E(X^3 - 2X + 15)$

6. Phương sai

$\text{var } X = DX = E[(X - EX)^2]$.

Phương sai là trung bình bình phương sai lệch giữa các giá trị X và $E(X)$

D = Dispension = độ phân tán = variance

TH1. X là BNN rời rạc

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

$\Rightarrow \text{var } X = (x_1 - EX)^2 p_1 + (x_2 - EX)^2 p_2 + \dots + (x_n - EX)^2 p_n$

TH2. X là BNN liên tục

$$\text{var } X = DX = \int_{-\infty}^{+\infty} (x - EX)^2 \cdot f(x)\,dx$$

* T/chất

① $\text{var}(c) = 0$; c hằng số

② $\text{var}(c.X) = c^2 . \text{var } X$

(34) $Var(X+Y) = VarX + VarY$

nếu X, Y độc lập

Vd1 cho X và Y là 2 BNN rời rạc, có bảng pp xác suất

x	-3	1	2	5
f	0,1	0,4	0,2	0,3

Y	-4	0	3	7
P	0,15	0,2	0,25	0,4

a. tìm VarX, VarY

b. $Var(X^2)$

c. $Var(Y^3 - 2X)$ phương sai k° có dấu trừ

6. các qui luật pp xác suất thường gặp

a. Luật Bécnuli

Ta nói biến ngẫu nhiên X có luật pp Becnuli, kí hiệu

$X \sim B(1, p)$

Nếu X có bảng pp

x	0	1
f	p	q

với $q = p - 1$;

khi đó $EX = p$

$VarX = pq$

+ cách kiểm chứng

$E(X^2-1) = \int (x^2-1) f(x)dx$

Độ lệch chuẩn $\sigma = \sqrt{Var X} = \sqrt{DX}$ Thứ.....Ngày.....Tháng.....Năm.....

+ có sự kiện A cần quan - { ... }

với $P(A) = p$

gọi X là số lần xuất hiện A trong 1 phép thử

$X \sim B[1, p]$

vd tung 1 cục xí ngầu gọi X là số lần xuất hiện một số chấm là bội số của 5. Luật pp cho X.

b. Luật nhị thức

ta nói X có luật pp nhị thức, kí hiệu $X \sim B(n,p)$ nếu

Binomial

+ ta có: phép thử = tung 1 cục xí ngầu

tung 1 cục xí ngầu + sự kiện A = { xuất hiện có số chấm là bội của 5 }

với $P(A) = \frac{1}{5}$

gọi X là số lần xuất hiện mặt số chấm là bội của 5 trong 1 lần tung cục xúc xắc

$\Rightarrow X \sim B(1, \frac{1}{6})$.

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

$\forall k = 0, 1, 2, \ldots$

X	0	1 ...	n
P	$C_n^0 p^0 q^{n-0}$		$C_n^n p^n q^{n-n}$

Phương pháp kiểm chứng $X \sim B(n,p)$

+ có phép thử

+ có sự kiện A cần quan tâm

$P(A) = p$

+ lặp lại phép thử n lần độc lập

+ gọi X là số lần xuất hiện A trong n phép thử

$X \sim B(n, p)$

Vd Một cái máy sx ra sp với tỉ lệ làm hỏng sp là 20%. Máy sản xuất ra 18 sp. Tính xs sau cho trong 18 sản phẩm có: + ta có mô hình nhị thức

a. có đúng 3 sp hỏng + phép thử: máy sx ra 1 sản phẩm

b. 6 sp hỏng + sự kiện A = { sp làm ra hỏng }

6 → 10 sp hỏng $P(A) = p = 0,2$

có số sp tốt ≥ 80% sp làm ra → số sp hỏng ≤ 3

+ lặp lại phép thử 18 lần

gọi X là số sp hỏng trong 18 sp làm ra

$\Rightarrow X \sim B(18, 0,2)$

a. $P(X=3)$

b. $P(X \le 6)$

$= P(6) + \dots P(6)$

c. Luật poisson: trường hợp riêng pp

Ta nói X có luật pp poisson kí hiệu $X \sim P(\lambda)$ nếu

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

Phương pháp kiểm chứng

+ có phép thử

+ có sự kiện A cần qtâm

với $P(A) = p$ (nhỏ)

+ lặp lại phép thử n lần

n rất lớn

+ gọi X là số lần thực hiện A phép thử (trong n phép thử)

Tuần ..... Ngày ..... Tháng ..... Năm .....

$P = P(A) = \frac{2}{1000}$ (nhỏ)

→ $X \sim P(\lambda)$

với $\lambda = np$

khi này $EX = varX = \lambda$

Vd 3 Một cái máy sx ra vaccan với tỉ lệ lớn và tỉ lệ làm hỏng sp là $2‰$. Máy sx ra 6000 lọ. tính sx ra

a. có đúng 10 lọ hỏng $(P(X=10)) = \frac{12000^{10}}{10!} \cdot e^{-12000}$

b. có từ 10% → 15% số lọ bị hỏng

c. có tối đa 5% số lọ làm ra bị hỏng

d. có ít nhất 96% số lọ làm ra là tốt

n đgl lớn ↔ $n > 30$

xác suất p nhỏ ↔ $p < 5\%$

$X \sim P(\lambda)$

$\lambda = np = 2$

$e^{-\lambda} \lambda^k / k!$

* Sửa bài tập :

câu 41.

Phép thử: ở một ống sợi

câu 33.

Ta có bảng phân phối cho 2 biến ngẫu nhiên X và Y ta nhân xác suất $P(X=1)$ và $P(Y=-2)$

Y \ X	1	2	3
-2	0,05	0,15	0,3
-1	0,03	0,09	0,18
0	0,02	0,06	0,12

ta có bảng xác suất

X+Y	-1	0	1	2	3
P					

có bao nhiêu giá trị mà làm cho X+Y bằng 0 ta cộng các xác suất tương ứng lại

* Luật pp chuẩn:

ta nói X là BNN tục có dạng pp chuẩn kí hiệu là $X \sim N(\mu, \sigma^2)$ nếu X có hàm mật độ

Normal

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

với

$$EX = \int_{-\infty}^{+\infty} x f(x) = \mu$$

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

khi $\mu = 0$ và $\sigma = 1$ thì ta nói X có dạng pp chuẩn tắc Normal standard khi đó X có hàm mật độ

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

$\phi(x)$ là hàm pp.

$\varphi(x)$ là hàm mật độ } phân phối chuẩn tắc.

Thứ.....Ngày.....Tháng.....Năm.....

Đồng thời hàm pp lúc này là.

$$\phi(x) = P(X \le x) = \int_{-\infty}^{x} g(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

* Khi $X \sim N(\mu, \sigma^2)$ thì ta có công thức xấp xỉ

$$P(a \le X \le b) = \phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)$$ μ giá trị trung bình

Khi $X \sim N(0,1)$

thì $P(a \le X \le b) = \phi(b) - \phi(a)$

trong đó

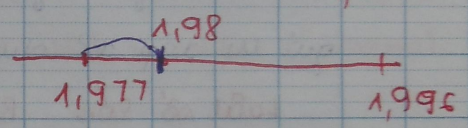
$\phi(x)$ giá trị của hàm phân phối của $X \sim N(0,1)$

ứng với giá trị x cần tính

tra bảng phụ lục 5. $\phi(1,98)$

vd. $\phi(1,960)$ 1,977 1,996.

$\phi(2,576)$



hay $\phi(c) = 0,9 \Rightarrow c =$

$\phi(c) = 0,940$.

$\phi(c) = 0,996$.

$\phi(-1,645) =$ lưu ý. $\phi(-c) = 1 - \phi(c)$.

$\phi(-1,881) =$

7. Các định lí giới hạn

a. Luật số lớn (LSL)

các biến NN $X_1, X_2, \ldots$ độc lập với nhau, đgl thỏa luật số lớn

$$\lim_{n \to \infty} P\left(\left|\frac{X_1 + \ldots + X_n}{n} - \frac{EX_1 + \ldots + EX_n}{n}\right| \le \varepsilon\right)$$

Cho nên, nếu $X_1, X_2, \dots X_n$ độc lập có cùng pp và có kì vọng hữu hạn

$$\mu = EX_1 = EX_2 = \dots = EX_n$$

Khi đó chúng thoả LSL nghĩa là

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+\dots+X_n}{n} - \mu\right| \le \varepsilon\right) = 1$$

Mức này, ta viết

$$\frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow{xs} \mu$$

ta gọi là hội tụ xác suất về μ

* Ứng dụng

Khi đo 1 đại lượng $\mu = EX$ nào đó khi ta lặp nhiều lần rồi lấy trung bình các giá trị đo ta lấy giá trị xấp xỉ cho μ

Luật số lớn Bernuli

xét mô hình như thế.

có phép thử

có sự kiện A cần quan tâm

$P(A) = p$

gọi X_i là số lần xuất hiện A ở phép thử thứ i

$X_i \sim B(1, p)$.

gọi $X = X_1 + X_2 + \dots + X_n$

thì X là số lần xuất hiện A trong n phép thử

Khi đó $X_1, X_2, \dots, X_n$ thoả LSL nghĩa là

$$\lim_{n \to +\infty} P\left(\left| \frac{x_1 + \dots + x_n}{n} - p \right| \leq \varepsilon \right) = 1$$

ta đặt

$f_n = \dfrac{x_1 + x_2 + \dots + x_n}{n}$ = tần suất xuất hiện A trong

n phép thử

ta viết lại là

$$f_n \xrightarrow{xs} p = P(A)$$

* Ứng dụng

Dùng để xác định xác suất của sk A trong phép thử đó n lần, độc lập nhau, rồi lấy tần suất của A làm giá trị cho xs của A

* Định lí giới hạn trung tâm: Moiver Laplace

xét mô hình như thế: (central Limited Theorem).

+ có phép thử.

+ có sự kiện A cần quan tâm

với $P(A) = p$

gọi x_i là số lần xuất hiện biến cố A ở lần thứ i

$\rightarrow x \sim B(1, p)$.

+ gọi $x = x_1 + x_2 + \dots + x_n$

= số lần xuất hiện biến cố A trong n phép thử

$\rightarrow x \sim (n, p)$

Khi đó ta với $x_1, x_2, \dots, x_n$ thoả mãn đi giới hạn trung tâm, có nghĩa là:

$$\lim_{n\to\infty} P\left(\frac{X-np}{\sqrt{npq}} \le x\right) = \phi(x) \quad (n \geq 30)$$

Lúc này

$$\begin{cases} E(X) = np \\ Var(X) = npq \end{cases} \quad (n \geq 30)$$

Đặt $Y = \dfrac{X-np}{\sqrt{npq}}$

thì ta nói Y là dạng thu gọn (dạng quy tâm của X)

Khi $n \to \infty$ (n có giá trị rất lớn) $(n \geq 30)$

$X \sim N(np, npq)$

$Y \sim N(0,1)$

* Công thức gần đúng:

$$\boxed{P(a \le X \le b) = \phi\left(\frac{b-np}{\sqrt{npq}}\right) - \phi\left(\frac{a-np}{\sqrt{npq}}\right)}$$

ra $\varphi(\blacksquare)$ trong bảng phụ lục 5

Định lí giới hạn địa phương.

(Local limited theorems).

xét mô hình như thức

+ có phép thử

+ có sự kiện A với $P(A) = p$

gọi X_i là số lần thực hiện phép thử A ở phép thử thứ i

ta có $X_i \sim B(1, p)$

+ gọi $X = X_1 + X_2 + \dots + X_n$ = hằng số số lần xuất hiện

(khi $a \le X \le b$ ⇒ đlí giới hạn toàn

$X = k$ ⇒ đlí gh địa phương.

Thứ.....Ngày.....Tháng.....Năm.....

biến cố A trong n phép thử

⇒ $X \sim B(n, p)$

Khi đó $X_1, X_2, \dots, X_n$ thoả đlí giới hạn địa phương có nghĩa là:

$$\lim_{n \to \infty} P(X = k) = \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(k-np)^2}{2npq}}$$

* công thức xấp xỉ:

$$\boxed{P(X = k) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(k-np)^2}{2npq}}}$$

Vd: Gieo 1 đồng xu cân đối, đồng chất 200 lần. Tính xs sao cho trong 200 lần gieo đồng xu:

a. có đúng 80 lần sấp

b. số lần sấp từ 100 ÷ 150

c. có ít nhất 40 lần ngửa

d. số lần sấp ≥ 70% số lần gieo đồng xu

giải:

Bsấp 4, 5, ~~34~~, 39, 68, 69, 107, 108, 110, 111, 112, 113, 116, 117, 126, 134, 136, 137, 140, 142, 142, 143, 145, 146, 148, 150, 155, 156, 159, 161, 184, 189, 190, 193 → 199, 203, 206.

* sửa bsấp: xét mà tính như thức

câu 34. Phép thử A: bắn 1 viên đạn

sự kiện: = {bắn trúng mục tiêu}

với p(A) = 0,8.

gọi X_i là số lần bắn trúng mục tiêu trong lần bắn thứ i

$X \sim B(100; 0,5)$

lặp lại phép thử 100 lần.

X là số lần bắn trúng trong 100 lần.

a. $P(75 \le X \le 90)$

Đề giải xấp xỉ tâm

ta có $\phi\left(\frac{90 - 100.0,8}{\sqrt{100.0,8.0,2}}\right) - \phi\left(\frac{75 - 100.0,8}{\sqrt{100.0,8.0,2}}\right)$

$= \phi(\cancel{\ }) = 0 - \phi(-1,25)$

b. nhiều ít hơn 75 lần bắn

câu 117: xét mô hình như thức

phép thử A. bắn 1 viên đạn

sự kiện A { đạn trúng mục tiêu }.

ta có $P(A) = p = 0,3$.

gọi X_i là số lần bắn trúng ở lần thứ i

$X_i \sim (1, 0,3)$

lặp lại phép thử $n = 500$ lần

$X \sim B(500; 0,3)$

a. AD đề giải xấp xỉ phương

ta có $P(X = 120) =$

$\frac{1}{\sqrt{500.0,3.0,7}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(500 - 150)^2}{2.500.0,7.0,3}}$

b. AD đề gửi hạn thân

$P(80 \le x \le 150)$.

câu 39.

câu 107. ta có mô hình pp poisson.

+ có phép thử vận chuyển 1 chai.

+ sự kiện A = { chai bị vỡ }

với $p(A) = 0,001$.

+ lặp lại phép thử 8000 lần.

gọi X là số lần thực hiện hiện cố A trong n phép thử.

ta có $X \sim P(\lambda)$ với $\lambda = np = 8$

ta có $P(X=5) = \dfrac{8^5 \cdot e^{-8}}{5!}$

ta có $P(X > 10) = 1 - P(X \le 10)$

$= 1 - \sum_{n=0}^{10} \dfrac{8^i e^{-8}}{n!}$

8. vector ngẫu nhiên: (BNN nhiều chiều).

một bộ (n) BNN đgl 1 bộ vector ngẫu nhiên (n) chiều

Nếu X_i là các BNN rời rạc ⇒ ta có bộ rời rạc

Nếu X_i là các BNN liên tục ⇒ ta có bộ vector là một bộ liên tục

1.1 Bảng pp xác suất đồng thời của (X,Y)

ta có.

ma trận m×n.

X \ Y	y_1	y_2	y_3 ...	y_n	tổng
x_1	p_{11}	p_{12}		p_{1n}	p_{1a} (XS theo dòng)
x_2	p_{21}	p_{22}		p_{2n}	p_{2a}
⋮					
x_m	p_{m1}	p_{22}		p_{mn}	p_{ma}
tổng	p_{a1}	p_{a2}		p_{an}	1

(XS theo cột).

trong đó.

(x_i, y_j) = giá trị của (X,Y) với $1 \le i \le m$, $1 \le j \le n$

$p_{ij} = P(X = x_i, Y = y_j)$ = xs để BNN X nhận giá trị x_i và Y nhận giá trị y_j

$p_{ia} = P(X = x_i) = p_{i1} + p_{i2} + \dots + p_{in}$

= xs để X nhận giá trị x_i

$p_{aj} = P(X = y_j) = p_{1j} + p_{2j} + \dots + p_{mj}$

= xs để Y nhận giá trị y_j

bảng pp lẻ của X và Y.

+ của X

X	x_1	x_2	x_m
P	p_{1a}	p_{2a}	p_{ma}

+ của Y

Mật độ đồng thời → Mật độ lề $\xrightarrow{\text{nguyên hàm}}$ pp lề

Thứ.....Ngày.....Tháng.....Năm.....

Y	y_1	y_2	.....	y_m
tổng xs của cột	$P_{\cdot 1}$	$P_{\cdot 2}$		$P_{\cdot m}$ (xs cột)

Hàm phân phối xs đồng thời

Hàm pp xác suất đồng thời của X và Y.

$F(x,y) = P(X \le x,\ Y \le y)$

* tính chất.

① $0 \le F(x,y) \le 1$.

② $\lim\limits_{x \to +\infty} F(x,y) = \lim\limits_{y} (y)$ = hàm pp lề của Y.

(y là tham số). giới hạn của hàm ñ biến

$\lim\limits_{y \to +\infty} F(x,y) = F_X(x)$ = hàm pp lề của X

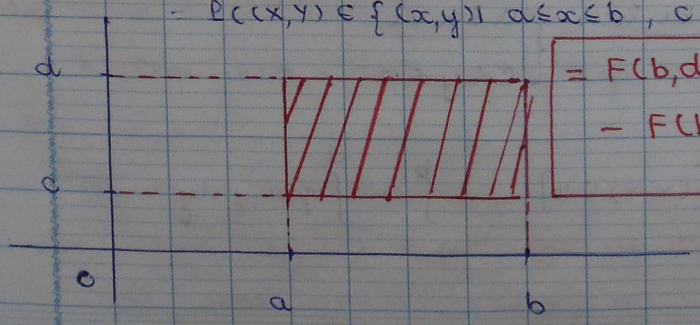
③ $\lim\limits_{x \to -\infty} F(x,y) = \lim\limits_{y \to -\infty} F(x,y) = \lim\limits_{\substack{x \to -\infty \\ y \to -\infty}} F(x,y) = \boxed{0}$.

một trong hai biến ∃ một biến tiến $-\infty$

④ $\lim\limits_{\substack{x \to +\infty \\ y \to +\infty}} F(x,y) = 1$.

⑤ $P(a \le X \le b,\ c \le Y \le d)$.

$= P((X,Y) \in \{(x,y) \mid a \le x \le b,\ c \le y \le d\})$.



$$\boxed{\begin{aligned} &= F(b,d) - F(a,d) \\ &- F(b,c) + F(a,c). \end{aligned}}$$

Nếu X, Y là rời rạc thì

$$F(x,y) = \sum_i \sum_j P_{ij}$$

Hàm mật độ phân phối đồng thời

Hàm mật độ đồng thời của (X, Y) liên tục là

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

* tính chất

① $f(x,y) \geq 0$.

② $$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,du\,dv$$

③ Hàm mật độ lề:

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

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

④ $$P[(X,Y) \in D] = \iint_D f(x,y)\,dx\,dy$$

Khi D là hình $\begin{cases} a \leq x \leq b \\ c \leq y \leq d \end{cases}$

$$\Rightarrow P[(X,Y) \in D] = \int_a^b dx \int_c^d f(x,y)\,dy$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dy\,dx = 1 \quad (*)$$

tìm hằng số a, 2.

Cho (X, Y) là vector ngẫu nhiên liên tục, có hàm mật độ đồng thời $f(x,y) = \begin{cases} 5x+3y & \text{khi } 0 \le x, y \le 2 \\ 0 & \text{nếu khác} \end{cases}$

a. Tìm $F(x,y)$.

b. Tìm $F_X(x,y)$, $F_Y(x,y)$

c. Tìm $f_X(x,y)$, $f_Y(x,y)$.

vd2. Cho (X, Y) là VTNN có bảng ppxs

xs trên

X \ Y	-3	2	4	
-1	4a	a	2a	= xs cột thứ nhất
0	3a	3a	4a	
5	3a	0a	5a	
				1.

a. Tìm a (ta có 25a = 1 ⇒ a = 1/25).

b. Tìm $F(x,y)$. (tính xs theo 16 miền con)

c. Tìm $F_X(x)$, $F_Y(y)$

vd2. a. ta có $25a = 1 \Rightarrow a = 1/25$

b. Tìm $F(x,y) = P(X \le x, Y \le y)$, $\forall x, y \in \mathbb{R}$.

TH1. $x < -1$.

→ $F(x,y) = 0$.

TH2. $-1 \le x < 0$ → có mấy giá trị làm cho $x < 0$?

. Khi $y < -3$, $F(x,y) = 0$.

. Khi $-3 \le y < 2$ → $F(x,y) =$

có mấy gtri làm cho $y < 2$?

QT BOOK

$F(x,y) = P(X=-1, Y=-3) = 4a$

. Khi $2 \le y < 4$

$F(x,y) = P(X=-1, Y=-3)$

$+ P(X=-1, Y=2) = 4a + a = 5a$

. Khi $y \geq 4$

$\Rightarrow F(x,y) = 4a + a + 2a = 7a$ = xác suất khi 1

TH3: * Khi $0 \le x < 5$ có mấy giá trị kèm cho $x < 5$

. Khi $y < -3$, $F(x,y) = 0$. $\left[\begin{array}{l} x = 0 \\ x = -1 \end{array}\right.$, $Y = 3$

. Khi $-3 \le y < 2$, $F(x,y) = 4a + 3a = 7a$

. Khi $2 \le y < 4$

$F(X=0, Y=3) + F(X=-1, Y=3) + F(X=-1, Y=2) + F(X=-1, Y$

$\rightarrow F(x,y) = 4a + 3a + a + 3a = 11a$

. Khi $y > 4 \Rightarrow F(x,y) = 11a + 2a + 4a = 17a$

TH4. Khi $x \geq 5$.

. Khi $y < -3$, $F(x,y) = 0$.

. Khi $-3 \le y < 2 \Rightarrow F(x,y) = 4a + 3a + 3a = 10a$

. Khi $2 \le y < 4 \rightarrow F(x,y) = 10a + 4a + 3a + 0 = 17a$

. Khi $y \geq 4 \Rightarrow F(x,y) = 25a$

Như vậy

$$F(x,y) \begin{cases} 0 & \text{khi } x < -1 \text{ hay } y < -3 \text{ cả 2 càng tốt} \\ 4a & \text{khi } -1 \le x < 0;\ -3 \le y < 2. \\ 5a & \text{khi } -1 \le x < 0;\ 2 \le y \le 4 \\ 7a & \text{khi } -1 \le x < 0.,\ y \geq 4 \\ & \text{hay } 0 \le x < 5,\ -3 \le y < 2. \\ 10a & \text{khi } x \geq 5,\ -3 \le y < 2 \\ 11a & \text{khi } 0 \le x < 5,\ 2 \le y < 4 \end{cases}$$

Mật độ $\xrightarrow[\text{nguyên hàm}]{\text{đạo hàm}}$ phân phối

Thứ.....Ngày.....Tháng.....Năm.....

14a khi $x \geq 5$; $k\ 2 < y < 4$.

17 a khi $0 \leq x < 5$, $y \geq 4$

1 khi $x \geq 5$, $y \geq 4$.

c. Phân phối lề của X.

Phân phối lề của Y

CẤU TRÚC ĐỀ THI

Câu 1: CTXS toàn phần + ct Bayes.

a. quét hết th xảy ra ở hộp 1 → cầu h đến hộp 2.

b.

Câu 2:

BNN rời rạc + liên tục

a. tìm λ.

b. tính EX, VarX

Câu 3:

B tập về luật pp poisson, nhị thức, luật chuẩn

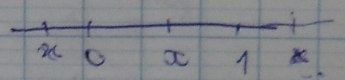
a.

b. có thể lập bảng ppxs: tính EX, VarX

Câu 15:

Cho X BNN $\sim U[0,1]$.

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



F(x) là hàm pp.

ta có $F(x) = \int_{-\infty}^{x} f(t)dt$

$-t\Big|_0^x = x$

Tiên tuế lễ, Hậu học văn QTBOOK

$$\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t)dt = 0 + \int_{0}^{x} 1\,dt$$

- Nếu $x<0$.

$$\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0\,dt = 0.$$

- Nếu $x>1$.

$$\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0 + \int_{0}^{1} 1\,dt + \int_{1}^{x} 0\,dt$$

$$= 1.$$

$$F(x) = \begin{cases} 0 & \text{———}. \\ x & \text{———}. \\ 1 & \text{———}. \end{cases}$$

Bài 18. Cho BNN X có dạng pp chuẩn tắc

$X \sim N(0,1)$

Đặt $Y = \left|\frac{X}{2}\right|$.

Tìm hàm mật độ của Y.

do X có dạng pp chuẩn tắc,

Hàm mật độ của X là

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

gọi $F(y)$, $f(y)$ lần lượt là hàm phân phối và hàm mật độ của BNN Y.

ta có $F(y) = P(Y \le y)$

$$= P\left(\left|\frac{X}{2}\right| \le y\right) \Rightarrow$$

Một trang vở một tương lai QI BOOK

$$= P\left(-\frac{X}{2} \le y \le \frac{X}{2}\right).$$

$= P(2y \le X \le -2y) = F(-2y) - F(2y).$

ta có: $Y = \left|\frac{X}{2}\right|$

tương ứng $y = \left|\frac{x}{2}\right|$.

mà

vậy hàm phân phối của BNN Y là.

$$F(y) = \begin{cases} 0 \text{ nếu } x < 0. \\ F(-2y) - F(2y) \end{cases}$$

Hàm mật độ của BNN Y.

$$f(y) = \begin{cases} 0 \text{ nếu } x < 0. \\ -2f(-2y) - 2f(-2y) = -2\left[f(-2y) + f(-2y)\right] \end{cases}$$

Mà X là BNN có dạng pp chuẩn tắc.

$X \sim N[0,1]$

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

$$\Rightarrow \varphi(-2y) = \frac{1}{\sqrt{2\pi}} e^{-(-2y)^2/2} = \frac{1}{\sqrt{2\pi}} e^{-2y^2}$$

$$\varphi(2y) = \frac{1}{\sqrt{2\pi}} e^{-(2y)^2/2} = \frac{1}{\sqrt{2\pi}} e^{-2y^2}$$

Vậy hàm mật độ của BNN Y là

$$\varphi(y) = \begin{cases} 0 \\ -4 \cdot \frac{1}{\sqrt{2\pi}} e^{-2y^2} \end{cases}$$

Thứ.....Ngày.....Tháng.....Năm.....

câu 2: tìm hàm mật độ $Y = X^2$.

ta gọi $F(x)$, $f(x)$

$F(y)$, $f(y)$ lần lượt là hàm phân phối và hàm mật độ của biến ngẫu nhiên X và Y

ta có:

~~$P(X) = P(X < x)$~~

~~$= P$~~

$$P(Y) = P(Y<y)$$
$$= P(X^2<y)$$
$$= P(-\sqrt{y} < X < \sqrt{y})$$
$$= F(\sqrt{y}) - F(-\sqrt{y})$$

vậy BNN Y có hàm phân phối là

$$F(y) = \begin{cases} 0 & \text{khi } x \le 0 \\ F(\sqrt{y}) - F(-\sqrt{y}) & \text{khi } x > 0. \end{cases}$$

ta có hàm mật độ.

$$f(y) = \begin{cases} 0 & \text{khi } x \le 0. \\ \frac{1}{2\sqrt{y}} f(\sqrt{y}) + \frac{1}{2\sqrt{y}} f(-\sqrt{y}). \end{cases}$$

Mà do BNN X có dạng phân phối chuẩn rồi

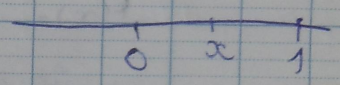
$\Rightarrow$ X có hàm mật độ là

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

$$\Rightarrow \varphi(y) = \begin{cases} 0 \text{ khi } y \le 0. \\ \frac{1}{\sqrt{2\pi y}} \left[e^{-y/2} \right] \end{cases}$$

Câu 16: Cho X có dạng pp đều. $X \sim U[0,1]$

do X có dạng pp đều.



$\Rightarrow f(x)$ có hàm mật độ là

$$f(x) = \begin{cases} 0 \text{ khi } x \notin [0,1] \\ 1 \text{ khi } x \notin [0,1] \end{cases}$$

gọi F(y) và f(y) lần lượt là ho

ta có.

F(x), f(x), F(y), f(y) lần lượt là hàm phân phối và hàm mật độ của BNN X và Y.

ta có. $F(x) = \int_{-\infty}^{x} f(x)\,dx$

. Nếu $x \in [0,1]$ ta có $\int_{-\infty}^{x} f(x)\,dx = \int_{-\infty}^{0} f(t)\,dt$

Tiên học lễ, Hậu học văn

QT BOOK

$$= \int_{-\infty} f(t)\,dt = \int_{0}^{x} 1\,dt = t\Big|_{0}^{x} = x$$

. Khi $x \notin [0,1]$

Nếu $x < 0$.

ta có $\int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0$.

Nếu $x > 1$

ta có $\int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{1} f(t)\,dt + \int_{1}^{x} f(t)\,dt$

$= \int_{0}^{1} 1\,dt = t\Big|_{0}^{1} = 1$.

Vậy ta có hàm phân phối

$$F(x) = \begin{cases} 0 & \text{khi } x < 0. \\ x & \text{khi } x \in [0,1] \\ 1 & \text{khi } x > 1. \end{cases}$$

ta có $Y = -5 \ln X$

$\Rightarrow \ln X = -\frac{Y}{5}$

$\Rightarrow X = e^{-Y/5}$

ta có $F(Y) = P(Y \le y)$

$= P(-5 \ln X \le y)$

$= P(X \geq e^{-y/5})$

$= 1 - P(X < e^{-y/5})$

$= 1 - F(e^{-y/5})$

Mà theo qui tắc đặt tương ứng.

$y = -5\ln x$

$Y = -5\ln X$

tương ứng : $y = -5\ln x$

$\Rightarrow x = e^{-y/5}$.

Mà $F(x) = x$ khi $x \in [0,1]$

$\Rightarrow F(e^{-y/5}) = e^{-y/5}$ khi $e^{-y/5} \in [0,1]$

Vậy $F(y) = \begin{cases} 1 - e^{-y/5} & \text{khi } y \geq 0 \\ 0 & \text{khi} \end{cases}$

- Bình luận

Tìm hàm phân phối của của hàm mật độ

$f(x) = \begin{cases} e^{-x} & \text{khi } x > 0 \\ 0 & \text{khi } x \leq 0. \end{cases}$

tìm $F(x)$, EX và VarX

ta có $F(x) = \int_{-\infty}^{x} f(t)dt$

. TH1 $x > 0$. ta có : $\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} f(t)dt + \int_{0}^{x} f(t)dt$

$= \int_{0}^{x} e^{-t}dt = -e^{-t}\Big|_{0}^{x} = -e^{-x} - (-e^{0})$

$= 1 - e^{-x}$

Đặt

TH2. $x < 0$; ta có $\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0\,dt = 0$

Tiên học lễ, Hậu học văn QT BOOK

d. Bảng phân phối có điều kiện. Hàm mật độ có điều kiện

TH1: X, Y rời rạc

ta có

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i ; Y = y_j)}{P(Y = y_j)}$$

$$= \frac{P_{ij}}{P_{oj}}$$ (xs của cột thứ j)

và

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i ; Y = y_j)}{P(X = x_i)} = \frac{P_{ij}}{P_{io}}$$

Nên ta có:

Bảng pp có điều kiện

$(X \mid Y = y_j)$	x_1	x_2 ...	x_m
P	$\frac{P_{1j}}{P_{oj}}$	$\frac{P_{2j}}{P_{oj}}$	$\frac{P_{mj}}{P_{oj}}$

và

$(Y \mid X = x_i)$	y_1	y_2 ...	y_n
	$\frac{P_{i1}}{P_{i\cdot}}$	$\frac{P_{i2}}{P_{i\cdot}}$	$\frac{P_{in}}{P_{i\cdot}}$

TH2: X, Y liên tục :

hàm mdô đtbiên

$$f_{(X|Y=y_0)}(x) = f(x|y_0) = \frac{f(x, y_0)}{f_Y(y_0)}$$

→ mật độ lề

$$f_{(Y|X=x_0)} = f(y|x_0) = \frac{f(y, x_0)}{f_X(x_0)}$$

VD1. Cho X, Y là hai biến ngẫu nhiên liên tục có hàm mật độ

$$f(x,y) = \begin{cases} 4 - x - y - xy & \text{nếu } 0 \le x, y \le 2 \\ 0 & \text{nếu khác} \end{cases}$$

a. Tìm $F(x,y)$

b. Tìm $F_X(x,y)$, $F_Y(x,y)$

c. Tìm $f_X(x,y)$, $f_Y(x,y)$

d. Tìm $F_{(X|Y=\frac{3}{2})}(x)$ và $F_{(Y|X=1/4)}(y)$

(→ hàm theo Y).

VD2. Cho X, Y là 2 biến ngẫu nhiên có bảng pp xác đồng thời

X \ Y	-1	0	4	8
-4	2a	a	5a	3a
1	4a	3a	a	6a
3	a	9a	2a	4a
5	0	7a	3a	9a

11 11a

14a

nhìn dòng làm cho x = 1, ứ mấy gtri của x = 1

a. Tìm a. ta có 50a = 1 11a ⇒ a = 1/50

b. Tìm $F(x,y)$

c. Lập bảng pp xs có đk của X khi $Y=4$

và lập bảng pp xs có điều kiện của Y khi $X=1$

2. Sự độc lập

TH1, X, Y là hai biến ngẫu nhiên rời rạc

X, Y độc lập nhau, nếu

$$P_{ij} = P(X=x_i, P=y_j)$$
$$= P(X=x_i) \cdot P(Y=y_j)$$
$$= P_{i.} \times P_{.j} \quad \forall i,j$$

TH2: X, Y là hai biến ngẫu nhiên liên tục

X và Y độc độc lập với nhau, nếu

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

Vd1 Cho X, Y là hai biến ngẫu nhiên rời rạc có bảng pp xs đồng thời

X \ Y	-1	2	4
-2	2a	3a	a
1	6a	2a	0
3	a	4a	a

a. Tìm a

b. Hỏi X và Y có độc lập với nhau không? vì sao

Vd2 Cho (X,Y) là vector ngẫu nhiên có hàm mật độ đồng thời

$$f(x,y) = \begin{cases} x^2y^2 - 2x - 2y & \text{nếu } 0 \le x, y \le 2 \\ 0 & \text{nơi khác} \end{cases}$$

lẻ theo $x \Rightarrow y$ chạy $\Rightarrow dy$

Thứ.....Ngày.....Tháng.....Năm.....

6. Hỏi X và Y có độc lập với nhau không

f. Kì vọng

TH1: X, Y là hai biến rời rạc

$$EX = \sum_{i=1}^{m} x_i p_{i\cdot}$$ ($p_{i\cdot}$ – xs của dòng)

($p_{\cdot j}$: xs của cột)

$$EY = \sum_{j=1}^{n} y_j p_{\cdot j}$$

TH2: X, Y liên tục

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

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

g. Phương sai:

TH1: X, Y rời rạc

$$Var X = \sum_{i=1}^{m} (x_i - EX)^2 p_{i\cdot}$$

$$Var Y = \sum_{j=1}^{n} (y_j - EY)^2 p_{\cdot j}$$

hay

$$Var X = E(X^2) - [EX]^2$$

$$= \sum_{i=1}^{m} (x_i^2 p_{i\cdot}) - \left(\sum_{i=1}^{m} x_i p_{i\cdot}\right)^2$$

$$Var Y = E(Y^2) - (EY)^2$$

$$= \sum_{j=1}^{n} y_j^2 p_{\cdot j}$$

TH2. X, Y rời BNN liên tục

$$\operatorname{var} X = \int_{-\infty}^{+\infty} (x - EX)^2 f_X(x)\,dx$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - EX)^2 \cdot f_X(x)\,dx$$

$$\operatorname{var} Y = \int_{-\infty}^{+\infty} (y - EY)^2 \cdot f_Y(y)\,dy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}$$

B tập:

Va1.

$$f(x,y) = \begin{cases} 4 - x - y - xy & \text{nếu } 0 \le x, y \le 2. \\ 0 & \text{nếu khác} \end{cases}$$

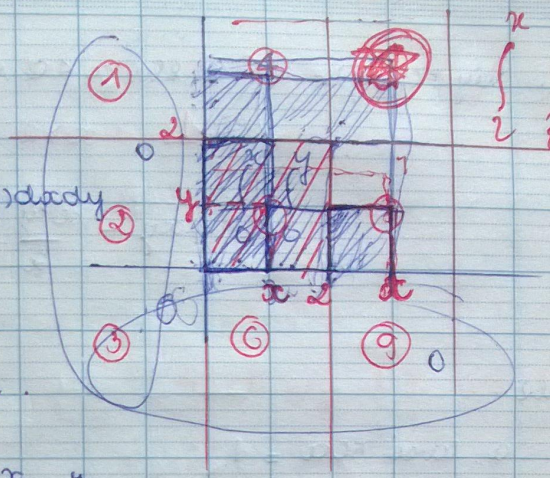
a. ta có $F(x,y) =$

$$F(x,y) \begin{cases} \int_{-\infty}^{x}\int_{-\infty}^{y} (4 - u - v - uv)\,du\,dv & \text{nếu } 0 \le u, v \le 2 \\ 0 & \text{nếu khác} \end{cases}$$

$$= \int_{-\infty}^{0}\int_{-\infty}^{0} + \int_{0}^{2}\int_{0}^{2} + \int_{2}^{x}\int_{2}^{y}$$

Thứ..... Ngày..... Tháng..... Năm.....

$$\int_{2}^{x}\int_{2}^{y} = 1.$$



* TH1: $x<0$ hay $y<0$.

ta có:

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\,dx\,dy$$

$$= 0.$$

* TH2: $0 \le x, y \le 2$.

ta có $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv$.

$$= \int_{0}^{x}\int_{0}^{y} (4-u-v-uv)\,du\,dv$$

$$= 4xy - \frac{x^2y}{2} - \frac{xy^2}{2} - \frac{x^2y^2}{4}$$

* TH3: $0 \le x \le 2$, $y \geq 2$.

ta có $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv$

$$= \int_{0}^{x}\int_{y}^{2} 0\,du\,dv = 0 \text{ khi } y \geq 2$$

$$F(x,y) = \int_{2}^{x}\int_{-\infty}^{y} 0\,du\,dv = 0 \text{ khi } x \geq 2.$$

$$F(x,y) = \int_{2}^{x}\int_{2}^{y} 0\,du\,dv = 0 \text{ khi } x,y \geq 2.$$

Tiên học lễ, Hậu học văn

QI BOOK

Vậy $F(x,y) = \begin{cases} 0 & \text{khi } x \le 0 \text{ hoặc } y \le 0 \text{ hoặc} \\ 1 & x \ge 2 \text{ hoặc } y \ge 2 \\ 4xy - \frac{x^2y}{2} - \frac{xy^2}{2} - \frac{x^2y^2}{4} \end{cases}$

Câu 2:

a. ta có $50a = 1$

b. * Khi $x < -4 \rightarrow F(x,y) = 0$.

* Khi $-4 \le x < 1$.

TH1. Nếu $y < -1$, ta có $F(x,y) = 0$.

Nếu $-1 \le y < 0$, ta có

$F(x,y) = 2a$.

TH2. Nếu $0 \le y \le 4$, ta có $F(x,y) =$

$2a + a = 3a$.

TH3. Nếu $4 \le y \le 8$. ta có

$F(x,y) = 2a + a + 4a = 8a$.

TH4. Nếu $y > 8$, ta có $F(x,y) = 11a$.

* Khi $1 \le x < 3$.

. Nếu $y < -1$ ta có $F(x,y) = 0$

. Nếu $-1 \le y < 0$ ta có $F(x,y) = 2a + 4a = 6a$

. Nếu $0 \le y < 4$ ta có $F(x,y) = (2 + 4 + 3 + 1)a =$

. Nếu $4 \le y < 8$ ta có $F(x,y) = (2 + 4 + 3 + 1 + 5 +$

$= 15a$

Thứ ..... Ngày ..... Tháng ..... Năm .....

(co = co operate).

k. Hiệp phương sai (covariance) → có gtri âm.

Hiệp phương sai của X và Y

kí hiệu $cov(X,Y) = E[(X-EX)(Y-EY)]$

TH1 X, Y rời rạc

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

TH2: X, Y liên tục tương tự EX và EY.

$$\boxed{cov(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-EX)(y-EY)\, f(x,y)\,dx\,dy}$$

* Tính chất:

(i) $cov(X,Y) = cov(Y,X)$

(ii) $cov(X+Y,Z) = cov(X,Z) + cov(Y,Z)$

(3i) $cov(kZ) = k.cov(Z)$

(4i) $cov(X,X) = varX$

(5i) $|cov(X,Y)| \le \sqrt{varX.varY}$

(bất cauchy - schwaste)

(6i) $[E(XY)]^2 \le E(X^2)E(Y^2)$

(i). Hệ số tương quan: (r - relation ship)

Hệ số tương quan của hai BNN X và Y

kí hiệu

$$\boxed{r_{X,Y} = \frac{cov(X,Y)}{\sqrt{varX.varY}}}$$

= 1: tương quan chặt. X ↑, Y ↑.

= -1 tương quan ngược

* tính chất:

(1i). $-1 \le r_{X,Y} \le 1$.

(2i) Nếu X, Y độc lập thì $r_{X,Y} = 0$

j. Kì vọng có điều kiện (có thể ra giá trị âm)

$$E(Y|X=x_0) = \int_{-\infty}^{+\infty} y\, f(y|x_0)\, dy$$

$$= \int_{-\infty}^{+\infty} y \cdot \frac{f(x_0, y)}{f_X(x_0)}\, dy$$

$$E(X|Y=y_0) = \int_{-\infty}^{+\infty} x \cdot f(x|y_0)\, dx = \int_{-\infty}^{+\infty} x \cdot \frac{f(x, y_0)}{f_Y(y_0)}\, dx$$

Và cho X, Y là vector ngẫu nhiên có hàm mật độ đồng thời

$$f(x,y) = \begin{cases} Axy & \text{khi } 0 \le x, y \le 1 \\ 0 & \text{nơi khác} \end{cases}$$

a. Tìm $F(x,y)$.

b. tìm $E(Y|X=\frac{1}{8})$ và $E(X|Y=\frac{3}{4})$

CHƯƠNG 3: ƯỚC LƯỢNG THAM SỐ

1. Một số khái niệm cơ bản.

. giả sử ta cần nghiên cứu đặc tính X trên các phần tử của tổng thể E → Total

- gọi N là số lượng các phần tử của E

vd1. X = chiều cao (cm)

E = tất cả các sinh viên của ĐHQG TpHCM

vd2. X = cân nặng

E = tất cả ~~trạm~~ cá ba sa thực khu vực ĐBSCL tháng 4/2003

. Thông thường, ta không thể quan sát tất cả các phần tử của E, do:

+ số lượng phử rất lớn

+ thời gian và kinh phí không cho phép

+ có thể làm hại các phần tử của E

→ do vậy, ta chỉ qsát 1 số phần tử, ghi nhận lại kq mà ta thường gọi là mẫu cỡ (n) rồi suy luận ra đặc tính của X ở một mức độ tin cậy nhất định

. Để có thể suy đoán đc chính xác, thì pp lấy mẫu phải thoả mãn:

+ các phử của mẫu phải được lấy ra chọn 1 cách ngẫu nhiên

+ các phử của E phải có đồng khả năng đc lấy ra làm mẫu

+ các phần tử của mẫu phải được lấy 1 cách độc lập

2. Các giá trị đặc trưng của mẫu.

a. trung bình mẫu - kì vọng mẫu

- giả sử ta cần nghiên cứu đặc tính X trên n phần tử.

+ gọi $x_1, x_2, \dots, x_n$ là (n) giá trị ta ghi nhận của X

TH1. Nếu mẫu có lặp

ta có bảng sau

$X = x_i$	n_i (số lần lặp)
x_1	n_1
x_2	n_2
...	...
x_k	n_k
tổng	n

trong đó n_i là số lần lặp lại của x_i

$$n_1 + n_2 + \dots + n_k = n$$

ta có.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{k} n_i x_i = \text{trung bình mẫu}$$

Vd. Đo chiều cao của 25 sv UIT ta thu được kq

164	170	180	172	168
162	158	164	170	164
162	155	180	171	167
166	162	158	155	164
158	164	172	168	165

Một trang vở một tương lai QT BOOK

Mới chiều cao thanh?

7+12. Mẫu chia khoảng

X	n_i	θ_i
$(a_1, a_2]$	n_1	θ_1
$(a_2, a_3]$	n_2	θ_2
⋮		
$(a_k, a_{k+1}]$	n_k	θ_k
tổng	n	

trong đó:

$$\theta_i = \frac{a_i + a_{i+1}}{2} = \text{điểm giữa } (a_i, a_{i+1}].$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{k} n_i \theta_i$$

* Phương sai mẫu:

Ví dụ: Để đo lường mức độ tiêu hao của nhiên liệu của 1 loại động cơ ô tô (đơn vị lít/1000 km) ta thí nghiệm và thu được kq.

X	n_i (số động cơ)
<4	4
4 - 5	6
5 - 6	10
6 - 7	7
7 - 8	5
>8	

s: độ lệch chuẩn có thực nghiệm

~~tb~~ ~~d/chủ~~

Thứ ..... Ngày ..... Tháng ..... Năm .....

s

Ước mức tiêu hao nhiên liệu tb của dòng cơ này

b. Phương sai mẫu:

TH1. mẫu có lặp:

$$s^2 = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})\,n_i$$: p/sai mẫu thực nghiệm

= p/sai mẫu thực ngh)

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{k}(x_i - \bar{x})\,n_i$$: phương sai mẫu có điều chỉnh

= p/sai mẫu có điều chỉnh

hoặc ta có thể dùng

$$s^2 = \overline{x^2} - (\bar{x})^2$$

với $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2$

hoặc ta có thể dùng bảng:

x	n_i	$n_i x_i$	$n_i x_i^2$
x_1	n_1	$n_1 x_1$	$n_1 x_1^2$
x_2	n_2	$n_2 x_2$	$n_2 x_2^2$
x_k	n_k	$x_k n_k$	$x_k x_k^2$
Tổng	n	T_1	T_2

DÙNG excel để tính các gtr

$$\bar{x} = \frac{T_1}{n}, \quad \overline{x^2} = \frac{T_2}{n}$$

$$s^2 = \overline{x^2} = (\bar{x})^2$$

$$S^2 = \frac{n}{n-1} s^2$$

TH2 Mẫu chia khoảng:

$$s^2 = \frac{1}{n} \sum_{i=1}^{k} n_i (\theta_i - \bar{x})^2$$

$$S^2 = \frac{n}{n-1} \sum_{i=1}^{k} n_i (\theta_i - \bar{x})^2$$

hoặc $s^2 = \overline{x^2} - (\bar{x})^2$

với $\overline{x^2} = \frac{1}{n} \sum_{i=1}^{k} n_i \theta_i^2$

$$\Rightarrow S^2 = \frac{n}{n-1} s^2$$

hoặc ta có thể dùng bảng:

x	θ	n_i	$\theta_i n_i$	$n_i \theta_i^2$
$(a_1, a_2]$	θ_1	n_1		
$(a_2, a_3]$	θ_2	n_2		
⋮				
$(a_k, a_{k+1}]$	θ_k	n_k	$n_k \theta_k$	$n_k \theta_k^2$
Tổng		n	T_1	T_2

Tiên học lễ, Hậu học văn QT BOOK

Mức độ phân tán của dữ liệu: s.

Thứ..... Ngày..... Tháng..... Năm.....

$$\bar{x} = \frac{T_1}{n}, \quad \overline{x^2} = \frac{T_2}{n}$$

$$s^2 = \overline{x^2} - (\bar{x})^2 \Rightarrow S^2 = \frac{n}{n-1} s^2$$

Vd. ta có bảng.

A	B	C	D	
x	n_i	θ_i	$n_i\theta_i$	$n_i\theta_i^2$
<4	4	3,5	14	49
4-5	6	4,5	27	121,5
5-6	10	5,5	55	302,5
6-7	7	6,5	45,5	295,75
7-8	5	7,5	37,5	281,25
>8	3	8,5	25,5	5527,125
tổng	35	36	204,5	1266,75

} mỗi khoảng cách đều 1 đơn vị

Mức tiêu hao năng lượng tb.

$$\bar{x} = \sum_{i=1}^{35}$$

ta có $\bar{x} = \frac{204,5}{35}$

3. ước lượng khoảng tin cậy (KTC) dành cho kì vọng

giả sử ta cần nghiên cứu đặc tính X trên n gần tử

. gọi $x_1, x_2, \dots, x_n$ là các giá trị đo được of X

gọi $\mu = EX$ là kì vọng mẫu trên toàn tổng thể

ta cần tìm μ_1 và μ_2 sao cho:

$$P(\mu_1 \le \mu \le \mu_2) = \gamma$$

với γ là mức độ tin cậy (ĐTC)

ta gọi $[\mu_1, \mu_2]$ là KTC ứng với ĐTC γ

TH1 $n \geq 30$, $\sigma_0^2 = \underline{var X}$ đã biết (đại lượng đã biết σ_0).

ta có KTC dành cho μ là

$$\boxed{\bar{x} - c.\frac{\sigma_0}{\sqrt{n}} \le \mu} \le \bar{x} + c.\frac{\sigma_0}{\sqrt{n}}$$

$$\mu_1 \le \mu \le \mu_2$$

với ĐTC $\gamma = 2\Phi(c) - 1$

$$\Rightarrow \Phi(c) = \frac{\gamma_1 + 1}{2} \Rightarrow c.$$

tra bảng pr. luc 5.

TH2 $n \geq 30$, $\sigma_0^2 = \underline{var X}$: chưa biết

ta có KTC dành cho μ là:

$$\bar{x} - c.\frac{s}{\sqrt{n}} < \mu \le \bar{x} + c.\frac{s}{\sqrt{n}} \qquad s^2 = \frac{n}{n-1} s_1^2$$

tự động tìm s: độ lệch chuẩn có điều chỉnh của mẫu đtra

với ĐTC: $\gamma = 2\gamma(c) - 1 \Rightarrow c = ?$

TH3 $n < 30$, $\sigma_0^2 = var X$ đã biết σ_0^2

tương ứng TH1

TH4 $n < 30$, $\sigma^2 = var X$ chưa biết

Mẫu cỡ nhỏ, phương sai chưa biết : phụ lục 7

KTC dành cho μ là

$$\bar{x} - c.\frac{S}{\sqrt{n}} \le \mu \le \bar{x} + c.\frac{S}{\sqrt{n}}$$

với ĐTC $= \gamma = 2F(c) - 1$

$\Rightarrow F(c) = \frac{1+\gamma}{2}$ (giá trị của cột)

$\Rightarrow c = ?$

Tra bảng phụ lục 7 ứng với cột có giá trị là $F(c)$, và dòng (ứng với bậc tự do) là $n-1$ ứng ra c.

ứng với dạng pp students

* Độ chính xác (ĐCX)

hay là bán kính của ước lượng

$$\varepsilon = \frac{1}{2}(\mu_2 - \mu_1)$$

$$= \frac{1}{2}\left[\bar{x} + c.\frac{\sigma_0}{\sqrt{n}} - \bar{x} + c.\frac{\sigma_0}{\sqrt{n}}\right)$$

$$= c.\frac{\sigma_0}{\sqrt{n}}$$

hoặc $\varepsilon = c.\frac{S}{\sqrt{n}}$ (σ_0 chưa biết)

④ cách ước lượng tỉ lệ bằng KTC.

gọi p là tỉ lệ các phần tử thoả đặc tính x cần nghiên cứu trên toàn bộ tổng thể

- ta cần tìm p_1 và p_2 sao cho:

$P(p_1 \le p \le p_2) = \gamma$

ta có KTC dành cho p là :

$$\hat{p} - c \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + c \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

với $\hat{p} = \frac{S_x}{n}$ với S_x là số lượng các phần tử thoả đặc

S_x K° phải độ lệch chuẩn có đc hình

tính x trên (n) phần tử khảo sát $\hat{p} = \frac{S_x}{n}$; K° 0.

và ĐTC γ $\gamma = 2\Phi(c) - 1) \Rightarrow c = ?$ K° đc giá trị âm

bằng cách tra bảng phụ lục (5).

* Bán kính của ước lượng. (ĐCX của UL)

$$\varepsilon = \frac{1}{2}(p_2 - p_1) = \frac{1}{2}\left[\hat{p} + c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} - \hat{p} + c\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

$$= c \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \rightarrow \frac{\varepsilon^2}{c^2} = \frac{\hat{p}(1-\hat{p})}{n}$$

$$\rightarrow n \ge \hat{p}(1-\hat{p}) \frac{c^2}{\varepsilon^2}$$

cho ε, từ đó tìm lại n,

ứng với : độ cx này cần bổ sung thêm ? nữa

giả sử ta cần nghiên cứu đặc trưng x trên toàn bộ tổng

thể G và ta gọi $\sigma^2 = Var X$ là phương sai của x trên G

ta cần tìm σ_1 và σ_2 sao cho

$P(\sigma_1 \le \sigma^2 \le \sigma_2) = \gamma$

với γ là ĐTC cho trước

ta cần tìm a và b sao cho : $P(a \le Y \le b) = \gamma$

trong đó

$$Y = \frac{\sum_{i=1}^{n} (x_i - \mu_0)^2}{a^2} \sim \chi^2(n)$$

là BNN có luật pp chi-bình phương với bậc tự do là n

. ta tìm a và b sao cho.

$$P(Y \le a) = \frac{1-\gamma}{2}$$

(tra bảng

$= \theta - \theta = 0.$

ta có $\bar{x}$ là ước lượng không chệch cho μ.

do $E(\bar{X}) = \mu$

S^2 là ước lượng không chệch

s^2 là ước lượng chệch cho $\sigma^2 = var\, x$

$\hat{p}$ là ước lượng không chệch dành cho p

tìm lấy giá trị $(x_1, x_2, \dots, x_n)$ ta tìm giá trị $\hat{\theta}$ là ước lượng xấp xỉ cho θ

. ta gọi $\hat{\theta}$ là ước lượng điểm cho θ

. sai số của ước lượng $\hat{\theta}$ cho θ là $\hat{\theta} - \theta$

5. Ước lượng KTC cho phương sai.

giả sử ta cần nghiên cứu đặc tính x trên toàn tổng thể E và ta gọi $\sigma^2 = Var X$ là phương sai trên E

+ ta tìm σ_1 và σ_2 sao cho:

$P(\sigma_1 \le \sigma^2 \le \sigma_1) = \gamma$ với γ là ĐTC cho trước.

TH1: $\mu = \mu_0 =$ EX đã biết (phương sai)

ta cần tìm hai số a và b sao cho: $(a \le Y \le b) = \gamma$

trong đó

$$Y = \sum_{i=1}^{n} \frac{(x_i - \mu_0)^2}{\sigma^2} = \chi^2(\mu)$$ là BNN số liệu

pp chi-bình phương với bậc tự do là n

ta tìm a và b sao cho:

$P(Y \le a) = \frac{1-\gamma}{2}$ (tra bảng phụ lục 6, ứng với cột $\frac{1-\gamma}{2}$ và dòng có giá trị là n) $\Rightarrow a = ?$

ta có $P(Y > b) = \frac{1-\gamma}{2} \Rightarrow P(Y \le b) = 1 - P(Y > b)$

$= \frac{1+\gamma}{2}$ tra bảng phụ lục 6, ứng với cột $\frac{1+\gamma}{2}$ và dòng có giá trị là $n \Rightarrow b = ?$).

ta có KTC dành cho σ^2 là $[\sigma_1, \sigma_2]$ với

$$\sigma_1 = \frac{\sum_{i=1}^{n} (x_i - \mu_0)^2}{b} \text{ và } \sigma_2 = \frac{\sum_{i=1}^{n} (x_i - \mu_0)^2}{a}.$$

TH2 $\mu = EX$ chưa biết ứng bậc tự do là $(n-1)$

$P(Y \le a) = \frac{1+\gamma}{2}$

ta tìm hai số a và b sao cho:

$P(a \le Y \le b) = \gamma$ với $\gamma = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

và BNN có luật pp chi-bình phương ứng với bậc tự do là $n-1$.

+ ta có: $P(Y \le b) = \frac{1+\gamma}{2}$ tra bảng phụ lục 6

ứng với cột có gtri là $\frac{1+\gamma}{2}$ và dòng ứng với bậc tự do là $(n-1)$ $\Rightarrow b = ?$

+ sau đó suy ra KTC do σ^2 như TH1

6. ước lượng điểm:

gọi θ là tham số đặc trưng X ta cần quan tâm đến toàn số tổng thể E

+ ở đây θ có thể là kì vọng, phương sai hay tỉ lệ. Giả sử giá trị $(x_1, x_2, \dots, x_n)$ ta tìm giá trị $\hat{\theta}$ là ước lượng xấp xỉ cho θ.

- Ta gọi $\hat{\theta}$ là ước lượng điểm cho θ
- sai số của ước lượng $\hat{\theta}$ cho θ là: $\hat{\theta} - \theta$
- $\hat{\theta}$ được gọi là ước lượng không chệch cho θ thì tb sai số của ước lượng là:

$E(\hat{\theta} - \theta) = E(\hat{\theta}) - E(\theta) = \theta - \theta = 0$

ta có.

Nếu đề bài k cho ĐTC ⇒ $\gamma = 95\%$.

⇒ mức ý nghĩa : 5%.

- $\bar{X}$ là ước lượng không chệch cho μ (do $E(\bar{X}) = \mu$
- S^2 là ước lượng không chệch cho $\sigma^2 = Var X$
- s^2 là ước lượng chệch cho $\sigma^2 = Var X$
- $\hat{p}$ là ước lượng không chệch cho p.

CHƯƠNG 4. KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ.

1. Một số khái niệm

- giả thiết thống kê : là một mệnh đề về tham số của BNN X mà ta quan tâm hoặc là mệnh đề về qui luật pp của X $\boxed{\alpha = 1 - \gamma}$. γ ĐTC trong chương 3.

Vd1 Khi kiểm định $\mu = EX$ thì ta có thể có các giả thiết

(i) $\begin{cases} \text{giả thiết } \mu = \mu_0 \\ \text{đối thiết (nghịch thiết) } K\ \mu > \mu_0 \end{cases}$

hoặc là:

(ii) $\begin{cases} \text{giả thiết } H = \mu = \mu_0 \\ \text{đối thiết } K : \mu < \mu_0 \end{cases}$ #ypothesis (gthiết)

hoặc là:

(iii) $\begin{cases} \text{giả thiết } \mu = \mu_0 \\ \text{đối thiết } \mu \neq \mu_0 \end{cases}$

Vd2. ta có giả thiết về luật pp của X như sau:

+ giả thiết H : X có luật pp poisson
+ đối thiết : X không có luật pp poisson

Đối với mỗi bài toán kiểm định, ta luôn có một mức độ sai lầm và ta gọi là mức ý nghĩa : α

2. Kiểm định giả thiết về kì vọng:

A. TH một mẫu.

Bài toán kiểm định là

$$\begin{cases} + \text{ giả thiết } H \quad \mu = \mu_0 \\ \text{Đối thiết } K: \quad \mu > \mu_0 \end{cases} \quad (i)$$

hoặc là

$$\begin{cases} \text{giả thiết } H. \quad \mu = \mu_0 \\ \text{Đối thiết } K. \quad \mu < \mu_0 \end{cases} \quad (2i)$$

hoặc là $\begin{cases} + \text{ giả thiết } H: \quad \mu = \mu_0. \\ + \text{ Đối thiết } K: \quad \mu \neq \mu_0 \end{cases}$ (3i)

ta tính đại lượng z như sau:

TH1. $n \geq 30$, $\sigma_0^2 = \text{var}X$ đã biết

$$z = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0} \quad (z \in \mathbb{R}).$$

TH2. $n \geq 30$, $\sigma^2 = \text{var}X$ chưa biết

$$z = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S}$$

TH3. $n < 30$, $\sigma_0^2 = \text{var}X$ đã biết giống TH1

TH4. $n < 30$, $\sigma^2 = \text{var}X$ ~~đã biết~~ chưa biết

như TH2

* Phép kiểm định tốt nhất là

$\begin{cases} + \text{ Bác chối } H \text{ nếu } z > c \quad (\text{tương ứng với TH 1 là dấu } >) \\ \text{và chấp nhận } H \text{ nếu } z \leq c \end{cases}$

(nếu ta xét (i))

(ứng với TH2 là <).

Hoặc là { bác bỏ H nếu $z < -c$
chấp nhận nếu $z \geq -c$ } nếu ta xét (2i).

Hoặc là:

{ + Bác bỏ H nếu $|z| > c$.
+ chấp nhận H nếu $|z| \leq c$ } nếu ta xét (3i).

Trong đó

c được tra bảng phụ lục (5) như sau:

– Nếu ta xét (1) và (2i) } khi tính z bằng TH1, TH2 và TH3.

$\phi(c) = 1 - \alpha$

(α) là mức ý nghĩa.....

– Nếu ta xét (3i) thì.

$$\boxed{\phi(c) = 1 - \frac{\alpha}{2}}$$

tuy nhiên khi tính z bằng (TH4) thì ta lại tra c từ bảng phụ lục (7) như sau.

$\boxed{F(c) = 1 - \alpha}$ [nếu ta xét (1) và (2i) tra dòng có giá trị là (n-1) và cột có giá trị $1-\alpha$ để tìm c)

và khi ta xét (3i) thì

$$\boxed{F(c) = 1 - \frac{\alpha}{2}}$$

(tra dòng (n-1) và cột là $1 - \frac{\alpha}{2}$ để tìm c)

B. Trường hợp 2 mẫu

B1. Hai mẫu cỡ lớn

giả sử ta cần nghiên cứu đặc tính x trên 2 tập hợp dân số (population)

+ tập hợp 1 có: n_1 : $\bar{x}_1$, s_1^2, μ_1 ← chưa biết

tập hợp 2 có: n_2, $\bar{x}_2$, $\bar{s}_2$, μ_2

chưa biết

ta cần so sánh μ_1 và μ_2

Bài toán kiểm định là

giả thiết H: $\mu_1 = \mu_2 = \mu_3$

đối thiết K: $\mu_1 > \mu_2$

$\mu_1 < \mu_2$

$\mu_1 \neq \mu_2$

tiếp theo, ta tính

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

nếu σ_1^2 = var x của tập hợp 1

σ_2^2 = var x của tập hợp 2

nếu σ_1, σ_2 chưa biết

hoặc

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} - \frac{\sigma_2^2}{n_2}}}$$

nếu σ_1^2 = var x của tập hợp 1

σ_2^2 = var x của tập hợp 2

đã biết

trong đó:

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{1}^{n_1} (x_1^{(1)} - \mu_0)^2$$

$$s_2^2 = \frac{1}{n_2 - 1} \sum_{n=1}^{n_2} (x_2^{(2)} - \mu_0)^2$$

nếu μ_0 đã biết

còn $x_i^{(1)}$ là tập hợp các giá trị x trên mẫu 1.

$x_i^{(2)}$ 2

. còn nếu μ_0 chưa biết thì ta tính s_1 và s_2 như sau

$$s_1^2 = \frac{1}{n_1 - 1} \sum (x_i^{(1)} - \bar{x}_1)^2.$$

$$s_2^2 = \frac{1}{n_2 - 1} \sum (x_i^{(2)} - \bar{x}_2)^2$$

Phép kiểm định tốt nhất là

+ Bác bỏ H nếu $z > c$.

+ Chấp nhận H nếu $z \leq c$ (nếu xét (i) và $\phi(c) = 1 - \alpha$)

hoặc

+ Bác bỏ H nếu $z < -c$.

+ Chấp nhận H nếu $z \geq -c$ (nếu xét (ii) và $\phi(c) = 1 - \alpha$)

hoặc là.

+ Bác bỏ H nếu $|z| > c$.

+ Chấp nhận H nếu $|z| \leq c$ (nếu xét (iii) và $\phi(c) = 1 - \frac{\alpha}{2}$)

Tất cả tra phụ lục (5) để tìm c.

B2. Hai mẫu cỡ nhỏ.

$n_1 \leq 30$, $n_2 < 30$.

. Bài toán kiểm định là (như B1)

nếu thì ta tính

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

với

$$(\hat{\sigma})^2 = \frac{(n_1-1).S_1^2 + (n_2-1).S_2^2}{n_1+n_2-2}$$

Bước kiểm định tốt nhất là

(như B1)

trong đó c sẽ được tra trong bảng phụ lục 5 nếu

$n_1+n_2-2 \geq 30$ (mẫu cỡ lớn).

ngược lại nếu $(n_1+n_2-2) < 30$ ta tra bảng phụ lục 7

mẫu cỡ bé

như sau:

ứng với dòng là (n_1+n_2-2) và cột là $(1-\alpha)$ nếu ta xét 1 hay 2i

và cột là $1-\frac{\alpha}{2}$ nếu ta xét 3i

3. Kiểm định giả thiết về tỉ lệ; chuẩn là mẫu cỡ lớ

A- Trường hợp 1 mẫu

Bài toán kiểm định là

+ giả thiết H: $p = p_0$

+ đối thiết k $p > p_0$ 1

$p < p_0$ 2i

$p \neq p_0$ 3i

tiếp theo ta tính

$$z = \frac{S_x - n.p_0}{\sqrt{n.p_0 q_0}}$$ (z có thể giá trị âm)

với S_x là số lượng phần tử thoả đặc tính x trên (n) phần tử khảo sát và $q_0 = 1 - p_0$.

phép kiểm định tốt nhất (như phần 2a)

(tra bảng phụ lục (5) để tìm c)

B- Trường hợp 2 mẫu :

p_1 là tỉ lệ các phần tử thoả đặc tính x trên tổng số n_1

p_2 n_2

Bài toán kiểm định là :

giả thiết $H = p_1 = p_2 = p$.

đối thiết K
$$\begin{cases} p_1 > p_2 & (1) \\ p_1 < p_2 & (2) \\ p_1 \neq p_2 & (3) \end{cases}$$

tiếp theo, ta tính z

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_0 \hat{q}_0 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ nếu biết p_0

hoặc $$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ nếu p_0 chưa biết

và $q_0 = 1 - p_0$

$\hat{p}_1 = \frac{S_x^{(1)}}{S_1}$, $\hat{p}_2 = \frac{S_x^{(2)}}{S_2}$

$S^1_{(x)}$ là số lượng phần tử thoả đặc tính x trên dân số 1

$S^2(x)$

$$\boxed{\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}} \quad , \quad \boxed{\hat{q} = 1 - \hat{p}}$$

Phép kiểm định tốt nhất là

(như phần 3a)

tương tương

tất cả dùng bảng phụ lục 5 để tìm con số c.

$$\boxed{\hat{p} = \frac{S_x^{(1)} + S_x^{(2)}}{n_1 + n_2}}$$

Sửa bài tập "KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ"

câu 10: ta có bài toán kiểm định là

+ giả thiết H: $p = p_0$

+ đối thiết K: $p \neq p_0$.

tiếp theo ta tính

$$Z = \frac{S_x - np}{\sqrt{n.p_0q_0}} = \frac{87 - 289.0,3}{\sqrt{289.0,3.0,7}}$$

$= 0,0385$

Phép kiểm định tốt nhất là.

ta tìm số c.

$$\phi(c) = 1 - \frac{\alpha}{2} = 1 - \frac{0,03}{2} = 0,985.$$

ra bảng phụ lục 5

ta có . $c = 2,170$.

Phép kiểm định tốt nhất là 17

từ chối nếu $|z| > c$

+ chấp nhận H nếu $|z| \leq c$

} → từ chối H.
lượng cartiscoit cao hơn
mức ý nghĩa $\alpha = 5\%$.

câu 1. ta có bài toán kiểm định là

+ giả thiết H $\mu = \mu_0$.

+ đối thiết K : $\mu < \mu_0$.

tiếp theo ta tính

$$z = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s^2} = \frac{\sqrt{20}(0,9-1)}{0,11} = -4,065.$$

$F(c) = 1 - \alpha = 1 - 0,01 = 0,99$.

ta tra bảng phụ lục 7, ứng với cột là 0,99 và bậc tự do là 19.

$\Rightarrow c = 2,539$.

Phép kiểm định tốt nhất là:

+ từ chối H nếu $z < -c$

+ chấp nhận nếu $z \geq -c$

} vậy ta từ chối H.

câu 2. ta có bài toán kiểm định là:

+ giả thiết H. $\mu_1 = \mu_2$.

+ Đối thiết K $\mu_1 < \mu_2$.

} tiếp theo ta tính.

$$X = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = 0,99$$: tra b

$\phi(c) = 1 - \alpha \Rightarrow \phi(c) = 1 - 0,01 = 0,99$.

tra bảng phụ lục số 5, ta có $c = 2,326$.

Phép kiểm

câu 4. ta có bài toán kiểm định là

+ giả thiết H: $p_1 = p_2$

+ đối thiết K: $p_1 \neq p_2$

xếp theo ta tính

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

ta có $\hat{p}_1 = \frac{S_x^{(1)}}{n_1} = \frac{14}{160} = 0,0875$.

$\hat{p}_2 = \frac{20}{225} = 0,0895$.

và $p = \frac{S_x^{(1)} + S_x^{(2)}}{n_1 + n_2} = \frac{14+20}{225+160} = 0,0883$

và $\Phi(c) = 1 - \frac{\alpha}{2} = 1 - \frac{0,05}{2} = 0,975$.

tra bảng phụ lục 5

ta có $c = 1,96$.

Phép kiểm định tốt nhất là,

câu 23. ta có bài toán kiểm định tỉ.

giả thiết H: $p_1 = p_2$.

đối thiết K: $p_1 \neq p_2$

ta tìm $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$

$\hat{p}_1 = \dfrac{S_1^{(x)}}{n_x} = \dfrac{20}{1000} = 0,2.$

$\hat{p}_2 = \dfrac{S_x^{(2)}}{n_2} = \dfrac{30}{900} = \dfrac{1}{30}.$

và $p = \dfrac{S_x^{(1)} + S_x^{(2)}}{n_1 + n_2} = \dfrac{50}{1000+900} = 0,0263.$

$\Rightarrow q = 1 - p = 1 - 0,0263 = 0,9737$

$\Rightarrow z =$

và $\Phi(c) = 1 - \dfrac{\alpha}{2}$

4. kiểm định giả thiết về phương sai:

4.1. Trường hợp 1 mẫu.

. bài toán kiểm định là

giả thiết H: $\sigma^2 = \sigma_0^2$

Đối thiết K: $\sigma^2 > \sigma_0^2$ (1)

$\sigma^2 < \sigma_0^2$ (2)

$\sigma^2 \neq \sigma_0^2$ (3)

Cấp thống kê tính

$$Z = \frac{(n-1).S^2}{\sigma_0^2}$$ (nếu μ_0 = EX chưa biết)

hoặc

$$Z = \frac{\sum (x_i - \mu_0)^2}{\sigma_0^2}$$ nếu μ_0 = EX đã biết

Miền tiêu chuẩn thích hợp là:

+ Bác bỏ H nếu $Z > \chi^2_{(n-1, 1-\alpha)}$

+ Chấp nhận H nếu $Z \leq \chi^2_{(n-1, 1-\alpha)}$

(nếu xét ① và μ = EX chưa biết)

hoặc

+ Bác bỏ H nếu $Z > \chi^2_{(n, 1-\alpha)}$

+ Chấp nhận H nếu $Z \leq$ ——

(nếu xét ② và μ_0 = EX chưa biết)

hoặc là

Bác bỏ H nếu $Z < \chi^2_{(n-1, 1-\alpha)}$

Chấp nhận H nếu $Z \geq$ ...

(nếu xét ③ và μ_0 chưa biết)

hoặc

+ Bác bỏ H nếu $Z < \chi^2_{(n, 1-\alpha)}$

+ Chấp nhận H nếu $Z \geq \chi^2_{(n, 1-\alpha)}$

(nếu xét ④ và μ_0 = EX đã biết)

hoặc là. bác bỏ H nếu $|z| > \chi^2_{(n-1, 1-\alpha)}$

chấp nhận H nếu $|z| \leq$

(nếu ta xét (1) và $\mu_0 = EX$ chưa biết)

hoặc

bác bỏ H nếu $|z| > \chi^2_{(n, 1-\alpha)}$

chấp nhận H nếu $|z| \leq$

Tra bảng phụ lục 6: phân vị χ^2 (khi bình phương)

4.2. Trường hợp 2 mẫu

ta có bài toán kiểm định là : $\sigma_1^2 = \sigma_2^2$

+ đối thiết K : $\sigma_1^2 > \sigma_2^2$

$\sigma_1^2 < \sigma_2^2$

$\sigma_1^2 \neq \sigma_2^2$

xét theo ta tính

$$z = \frac{S_1^2}{S_2^2}$$

ta có phép kiểm định tốt nhất là.

+ bác bỏ H nếu $z > \chi^2_{(n_1+n_2-2,\ 1-\alpha)}$

+ chấp nhận $\leq$

(nếu xét (1))

hoặc + bác bỏ H nếu $z < \chi^2_{(n_1+n_2-2,\ 1-\alpha)}$

+ chấp nhận H nếu $\geq$

nếu ta xét (2)

hoặc

+ bác bỏ H nếu $|z| > \chi^2(n_1+n_2-2, 1-\frac{\alpha}{2})$

+ chấp nhận $\leq$

(nếu 31).

5. kiểm định giả thiết về qui luật phân phối (kiểm định giả thiết fit hóp)

+ giả sử ta cần nghiên cứu đặc trưng X trên toàn bộ tổng thể E

Ta có bài toán kiểm định là

+ giả thiết H: X có qui luật pp X

+ đối thiết K: X không có qui luật pp X

. Ta lập bảng

Layer

tần số \ lớp	L_1	L_2		L_k	tổng
tần số thực nghiệm	N_1	N_2	...	N_k	n
tần số lí thuyết	N'_1	N'_2	...	N'_k	n

trong đó N_i là tần số thực nghiệm

= số lượng phần tử rơi vào lớp L_i

N'_i = tần số lí thuyết = số lượng phần tử của mẫu rơi vào lớp L_i nếu H đúng

ta có ~~$N'_i = n_i p_i$~~ $\boxed{N'_i = n_i p_i}$

với $p_i = P(X \in L_i \mid H \text{ đúng})$

= xác suất để X rơi vào lớp L_i nếu H đúng

(có thể đề cho sẵn hoặc ta phải tự tìm ra p_i)

sau đó, ta tính

$$Z = \sum_{i=1}^{k} \frac{(N_i - N'_i)^2}{N'_i} > 0 \quad \text{KO thể âm.}$$

ta có phép kiểm định tốt nhất là:

+ bác bỏ H nếu $Z > \chi^2_{(k-r-1,\ 1-\alpha)}$

+ chấp nhận H nếu $Z \leq$

k = số lớp dữ liệu = số cột

r = số lượng tham số ta cần ước lượng khi tính các p_i

bc. * luật pp chuẩn

b16. luật pp đều

b56.

gthiết H: X có qui luật pp poisson

Dthiết K: X KO có qui luật pp poisson

$X \sim P(\lambda)$

chưa biết.

dùng ước lượng $\bar{X} = EX$ để ước lượng

do ($\lambda = EX = VarX = \bar{X}$).

$\bar{X}$ đã được

các xs đưng đlí giới hạn địa phương

và ước lượng 1 tham số: $\boxed{k = 1}$

ứng với poison

- pp chuẩn ước lượng 2 tham số
- pp đều $X \sim U[a, b]$.

, k° ước lượng, $k = 0$.

các xs = nhau.

6. Kiểm định về sự độc lập

Trong thực tế, đôi khi ta cần kiểm định xem các tính chất nào nền tổng thể G có liên quan với nhau hay không?

Ta có bài toán kiểm định là

giả thiết H: "đặc tính A và B là độc lập nhau

đối thiết K: k° độc lập nhau

Tiếp theo ta lập bảng

đặc tính A \ đặc tính B	B_1	B_2	....	B_n	tổng
A_1	N_{11} / N'_{11}	N_{12} / N'_{12}			N_{A_1}
A_2	N_{21} / N'_{21}				N_{A_2}
⋮					
A_m	N_{m1} / N'_{m1}				N_{A_m}
	N_{B1}	N_{B2}		N_{Bn}	n

với $N'_{ij} = \dfrac{N_{Ai} \cdot N_{Bj}}{n}$

sau đó ta tính:

$$Z = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{(N_{ij} - N'_{ij})^2}{N'_{ij}}$$

Phép kiểm định tốt nhất là:

từ chối H nếu $Z \geq \chi^2_{(k-1)(c-1),\ 1-\alpha}$

chấp nhận H nếu $\leq$

k là số hàng, c là số cột

7 Kiểm định giả thiết về sự đồng nhất

giả sử có vì mẫu cùng quan tâm đến các yếu tố A, B, C, D

- ta cần đánh giá xem các mẫu này có đồng nhất với nhau hay không?

ta có bài toán kiểm định là.

giả thiết H: các mẫu là đồng nhất với nhau

đối thiết K: các mẫu không đồng nhất với nhau

xếp theo ta lập bảng.

yếu tố \ Mẫu	1	2	...	l	tổng
A	N_{11} N'_{11}	N_{12} N'_{12}			N_A
B	N_{21} N'_{21}	N_{22} N'_{22}			N_B
...	...	...			...
k	N_{k1} N'_{k1}	N_{k2} N'_{k2}			N_k
tổng	N_1	N_2	...	N_l	n

Hãy chia sẻ niềm vui, hãy viết ngày nên nghĩa OFBOOK

nếu theo ta tính.

$$Z = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{(N_{ij} - N'_{ij})^2}{N'_{ij}}$$

ta có phép kiểm định tốt nhất là:

bác bỏ H nếu $Z > \chi^2_{(k-1)(n-1),\, 1-\alpha}$.

chấp nhận H nếu $Z < \chi^2_{(k-1)(n-1),\, 1-\alpha}$

k số hàng

n : số cột của bảng.

tra bảng phụ lục 6, phân phối chi - bình phương

ĐH Quốc Gia HCM
Trường ĐH CNTT
Bộ môn Toán - Lý

Đề thi môn: **XÁC SUẤT THỐNG KÊ**
Thời gian: **90** phút
(Không sử dụng tài liệu, trừ bảng phụ lục)
Ngày thi: 05 / 06 / 2012
Lớp: **STA01.C21, STA01.C22**

Câu 1: (2 điểm)

Một đài dự báo khí tượng thủy văn muốn xem xét khả năng dự báo của mình nên đã tiến hành tổng hợp dữ liệu đã được thu thập và lưu trữ từ trước đây cho đến hiện tại. Nhà đài nhận thấy rằng: xác suất dự báo có nắng trong ngày không mưa là 80%, có nắng trong ngày mưa là 35%; đồng thời xác suất một ngày không có mưa là 75%. Chọn ngẫu nhiên một ngày sắp tới để dự báo. Hãy tính xác suất sao cho:

a/ Ngày này sẽ có nắng.

b/ Nếu ngày dự báo là có nắng thì hãy tính xác suất sao cho ngày này không có mưa.

Câu 2: (2 điểm)

Gọi X là biến ngẫu nhiên thể hiện cho khối lượng của một loại sản phẩm do nhà máy ABC sản xuất ra (đơn vị tính là gram). Giả sử rằng $X \sim N(100;16)$. Sản phẩm được xem là đạt tiêu chuẩn nếu khối lượng của nó đạt từ 93 g đến 107 g.

a/ Hãy tìm tỷ lệ phế phẩm của nhà máy này.

b/ Nếu nhà máy sản xuất ra 500 sản phẩm, hãy tính xác suất sao cho có trên 85% sản phẩm đạt tiêu chuẩn.

Câu 3: (2 điểm)

Cho vector ngẫu nhiên (X,Y) có hàm mật độ đồng thời

$$f(x,y)=\begin{cases} Ax^2y & khi\ 0\le x\le 1; 0\le y\le 2 \\ 0 & \text{ở những nơi khác} \end{cases}$$

a/ Hãy xác định hằng số A.

b/ Hỏi X và Y có độc lập nhau hay không? Vì sao?

Câu 4: (2 điểm)

Để nghiên cứu tuổi thọ của một loại thiết bị (đơn vị tính là tháng), người ta tiến hành kiểm tra ngẫu nhiên 16 thiết bị thì thu được kết quả sau: 98 102 96 98 100 104 110 102 106 90 107 115 98 114 100 102. Giả sử rằng tuổi thọ của thiết bị là biến ngẫu nhiên X có phân phối chuẩn.

a/ Hãy tìm KTC 95% dành cho tuổi thọ trung bình của thiết bị.

b/ Nếu muốn độ tin cậy là 99% và độ chính xác là 3 tháng thì cần phải điều tra thêm ít nhất bao nhiêu thiết bị nữa?

Câu 5: (2 điểm)

Một cái máy sản xuất ra sản phẩm. Nếu dùng phương pháp A thì trong 400 sản phẩm làm ra có 16 phế phẩm. Nếu dùng phương pháp B thì trong 300 sản phẩm làm ra có 15 phế phẩm. Như vậy, với mức ý nghĩa $\alpha = 0,01$ hãy xét xem hiệu quả của 2 phương pháp trên có giống nhau hay không?

HẾT

ĐH Quốc Gia HCM
Trường ĐH CNTT
Bộ môn Toán – Lý

Đề thi môn: **XÁC SUẤT THỐNG KÊ**
Thời gian: **90** phút
(Không sử dụng tài liệu, trừ bảng phụ lục)
Ngày thi: 28 / 08 / 2012
Lớp: **STA01.C31**

Câu 1: (2 điểm)

Một nhân viên quảng cáo tiến hành nghiên cứu sở thích xem TV của những người đã lập gia đình tại một vùng dân cư. Từ số liệu thu thập được, anh ta thấy rằng: có 65% các ông chồng thích xem TV. Khi ông chồng thích xem TV thì có 45% các bà vợ cũng thích xem TV; còn khi chồng không thích xem TV thì chỉ có 28% các bà vợ thích xem TV. Hãy tìm xác suất sao cho:

a/ Người vợ thích xem TV.

b/ Nếu vợ thích xem TV thì chồng cũng thích xem TV.

Câu 2: (2 điểm)

Một loại chi tiết máy được xem là đạt yêu cầu kỹ thuật nếu trị tuyệt đối sai lệch giữa đường kính của nó so với đường kính thiết kế không quá 0,33 mm. Biết rằng đường kính của trục máy là BNN có phân phối chuẩn $X \sim N(\mu_0, \sigma_0^2)$, trong đó $\sigma_0 = 0{,}3$ mm, và đường kính thiết kế $= \mu_0 + 0{,}321$ mm.

a/ Hãy tìm tỷ lệ chi tiết máy đạt yêu cầu kỹ thuật.

b/ Hãy tìm xác suất sao cho trong 200 chi tiết máy loại này có ít nhất 80% chi tiết đạt yêu cầu kỹ thuật.

Câu 3: (2 điểm)

Cho biến ngẫu nhiên X (đơn vị tính là tháng), là tuổi thọ của một loại thiết bị, có hàm mật độ xác suất

$$f(x) = \begin{cases} Axe^{-x/2} & khi \quad x \geq 0 \\ 0 & khi \quad x < 0 \end{cases}$$

a/ Hãy xác định hằng số A.

b/ Hãy tìm hàm phân phối xác suất của X; từ đó tính xác suất sao cho thiết bị có tuổi thọ từ 6 tháng đến 1 năm.

Câu 4: (2 điểm)

Để xác định tỷ lệ mắc bệnh B trong một vùng dân cư, người ta khảo sát ngẫu nhiên n người dân trong vùng đó, và tính ra tần suất mắc bệnh B là 0,32.

a/ Hãy tìm khoảng tin cậy cho tỷ lệ mắc bệnh B trong vùng dân cư này, với độ tin cậy 95% và 99%, ứng với $n = 625$.

b/ Hỏi cỡ mẫu là bao nhiêu để có độ chính xác là 0,03 (ứng với ĐTC 95% và 99%)?

Câu 5: (2 điểm)

Một cái máy sản xuất ra sản phẩm. Nếu dùng nguyên liệu A thì trong 800 sản phẩm làm ra có 320 sản phẩm loại 1. Nếu dùng nguyên liệu B thì trong 500 sản phẩm làm ra có 190 sản phẩm loại 1. Như vậy, với mức ý nghĩa $\alpha = 0{,}01$ hãy xét xem tác dụng của 2 nguyên liệu trên có giống nhau hay không?

HẾT

ĐH Quốc Gia HCM
Trường ĐH CNTT
Bộ môn Toán - Lý

Đề thi môn: **XÁC SUẤT THỐNG KÊ**
Thời gian: **90** phút
(Không sử dụng tài liệu)
Ngày thi: 15 / 01 / 2011
Lớp: **STA01.B14**

Câu 1: (2 điểm)

Một hộp gồm có 12 viên phấn đỏ và 8 viên phấn trắng. Lấy 1 viên phấn ra khỏi hộp rồi bỏ vào 1 viên phấn khác màu với nó. Sau đó lại lấy ra 1 viên phấn nữa. Tính xác suất để:

a/ Viên phấn lấy ra lần sau có màu trắng.

b/ Hai viên phấn lấy ra cùng màu.

Câu 2: (2 điểm)

Khi tiêm truyền một loại vaccine, người ta thấy trung bình có 1 trường hợp bị phản ứng trên 2000 trường hợp. Ta tiến hành tiêm cho 6000 người. Tính xác suất để:

a/ Có 3 trường hợp phản ứng.

b/ Nhiều nhất 4 trường hợp phản ứng.

Câu 3: (2 điểm)

Cho hàm số $f(x) = \begin{cases} \lambda x^2(2-x) & \text{nếu} \quad x \in [0,2] \\ 0 & \text{nếu} \quad x \notin [0,2] \end{cases}$

a/ Hãy xác định hằng số λ để $f(x)$ là hàm mật độ xác suất của một biến ngẫu nhiên X nào đó.

b/ Với giá trị λ tìm được ở câu a/, hãy tính kỳ vọng EX và phương sai $VarX$.

Câu 4: (2 điểm)

Để ước lượng số cá trong hồ, người ta đánh bắt 2000 con, đánh dấu rồi thả xuống hồ. Sau đó, người ta đánh bắt lên 460 con thì thấy có 92 con cá được đánh dấu. Với độ tin cậy 90% và 95%, hỏi số cá trong hồ có khoảng bao nhiêu con?

Câu 5: (2 điểm)

Tuổi thọ của bóng đèn X là đại lượng ngẫu nhiên có phân phối chuẩn, với giá trị trung bình là $\mu(X) = 2200$ giờ, và độ lệch tiêu chuẩn $\sigma(X) = 15$ giờ. Nghi ngờ tuổi thọ của bóng đèn đã thay đổi, người ta thắp thử 25 bóng thì thấy tuổi thọ trung bình của 25 bóng là $\overline{X} = 2190$ giờ. Với mức ý nghĩa $\alpha = 5\%$ và $\alpha = 1\%$, hãy kết luận về nghi ngờ nói trên.

HẾT

ĐH Quốc Gia HCM
Trường ĐH CNTT
Bộ môn Toán - Lý

Đề thi môn: **XÁC SUẤT THỐNG KÊ**
Thời gian: **90** phút
(Không sử dụng tài liệu, trừ bảng phụ lục)
Ngày thi: 17 / 06 / 2011
Lớp: **STA01.B21, STA01.B22**

Câu 1: (2 điểm)

Một người có: 8 con gà mái và 4 con gà trống, nhốt chung trong một lồng. Một người khác đến mua gà. Người bán gà bắt ngẫu nhiên 1 con gà. Người mua chấp nhận mua con đó.

a/ Tính xác suất để người đó mua con gà mái.

b/ Người thứ hai đến mua, người bán gà lại bắt ngẫu nhiên ra 1 con gà. Tính xác suất để bắt được gà trống, nếu người bán gà quên mất rằng con gà đã bán cho người thứ nhất là một con gà trống hay gà mái.

Câu 2: (2 điểm)

Xác suất không nảy mầm của hạt thóc giống là 0,004. Hãy tính xác suất sao cho khi chọn 2000 hạt thóc giống, ta có:

a/ Không ít hơn 12 hạt không nảy mầm.

b/ Số lượng hạt thóc nảy mầm phải lớn hơn 80%.

Câu 3: (2 điểm)

Cho (X,Y) là vector ngẫu nhiên có hàm mật độ:

$$f(x,y)=\frac{A}{1+x^2+y^2+x^2y^2}$$

a/ Hãy xác định hằng số A.

b/ Chứng minh rằng X và Y độc lập nhau.

Câu 4: (2 điểm)

Để ước lượng số chim trong một khu rừng, người ta dùng bẫy để bắt. Biết rằng lần đầu tiên có 800 con chim bị bắt, sau đó chúng được đánh dấu, rồi thả lại vào rừng. Sau đó, người ta bắt tiếp 140 con chim thì thấy có 28 con chim đã được đánh dấu. Với độ tin cậy 92% và 95%, hỏi số lượng chim trong khu rừng có khoảng bao nhiêu con?

Câu 5: (2 điểm)

Trong 82 người dùng café, có 34 người bị mất ngủ. Trong 96 người không dùng café có 20 người bị mất ngủ. Với mức ý nghĩa $\alpha = 0{,}05$, hãy xét xem café có gây mất ngủ hay không? (nghĩa là xét xem sự mất ngủ có phụ thuộc vào việc dùng café hay không?)

HẾT

ĐH Quốc Gia HCM
Trường ĐH CNTT
Bộ môn Toán - Lý

Đề thi môn: **XÁC SUẤT THỐNG KÊ**
Thời gian: **90** phút
(Không sử dụng tài liệu, trừ bảng phụ lục)
Ngày thi: 31 / 08 / 2011
Lớp: **STA01.B31**

Câu 1: (2 điểm)

Tỷ lệ người nghiện thuốc lá ở một vùng là 36%. Biết rằng tỷ lệ người bị viêm họng trong số những người nghiện thuốc là 75%, còn tỉ lệ người bị viêm họng trong số những người không nghiện thuốc là 28%.

a/ Chọn ngẫu nhiên một người để khám bệnh thì thấy rằng người ấy bị viêm họng. Tính xác suất người ấy nghiện thuốc.

b/ Nếu người đó không bị viêm họng, tính xác suất người đó nghiện thuốc.

Câu 2: (2 điểm)

Một xe tải vận chuyển 10000 chai rượu vào kho. Xác suất mỗi chai rượu bị vỡ là 0,002 khi vận chuyển. Hãy tính xác suất sao cho mỗi khi vận chuyển:

a/ Có không quá 10 chai bị vỡ.

b/ Có số chai không bị vỡ tối thiểu là 90%.

Câu 3: (2 điểm)

Cho vector ngẫu nhiên (X,Y) có hàm phân phối đồng thời

$$F(x,y)=\begin{cases}1-3^{-x}-3^{-y}+3^{-x-y} & khi \quad x\ge 0; y\ge 0\\ 0 & \text{ở những nơi khác}\end{cases}$$

a/ Hãy tìm hàm mật độ đồng thời $f(x,y)$.

b/ Hỏi X,Y có độc lập nhau hay không?

Câu 4: (2 điểm)

Một kho hàng có 500000 lọ vaccine. Người ta kiểm tra ngẫu nhiên 120 lọ thì thấy có 18 lọ bị hỏng. Như vậy, với độ tin cậy 92% và 99%, hãy xác định xem trong kho có khoảng bao nhiêu lọ vaccine bị hỏng?

Câu 5: (2 điểm)

Trong 200 bệnh nhận điều trị bệnh L bằng thuốc A thì thấy có 160 người khỏi, 30 người giảm bệnh và 10 người bị biến chứng. Còn trong 300 người điều trị bệnh L bằng thuốc B thì thấy có 18 người bị biến chứng, 216 người khỏi bệnh và 66 người giảm bệnh. Như vậy, với mức ý nghĩa $\alpha = 0{,}01$ hãy xét xem mức độ công hiệu của 2 loại thuốc trên có giống nhau hay không?

HẾT

ĐH Quốc Gia HCM
Trường ĐH CNTT
Bộ môn Toán – Lý

Đề thi môn: **XÁC SUẤT THỐNG KÊ**
Thời gian: **90** phút
(Không sử dụng tài liệu, trừ bảng phụ lục)
Ngày thi: 26 / 12 / 2011
Lớp: **STA01.C13, STA01.C14**

Câu 1: (2 điểm)

Hộp A có 10 bóng đèn màu đỏ, 4 bóng màu xanh, 6 bóng màu vàng. Hộp B có 8 bóng đèn màu đỏ, 7 bóng màu xanh, 4 bóng màu vàng. Từ hộp A ta lấy ra ngẫu nhiên 1 bóng đèn bỏ sang hộp B. Sau đó, từ hộp B ta lấy ra ngẫu nhiên 1 bóng đèn để xem màu. Hãy tính xác suất sao cho:

a/ Bóng đèn lấy ra ở lần thứ hai không phải màu xanh.

b/ Bóng đèn lấy ra ở lần thứ nhất và lần thứ hai khác màu với nhau.

Câu 2: (2 điểm)

Tỷ lệ mắc bệnh bẩm sinh trong dân số ở một địa phương là 0,001. Bệnh này cần sự chăm sóc đặc biệt lúc mới sinh. Biết rằng trong tháng 11 năm 2011 có 4000 trẻ được sinh ra tại địa phương này. Tính xác suất sao cho:

a/ Có không quá 8 trẻ cần sự chăm sóc đặc biệt.

b/ Có ít nhất 80% trẻ không cần sự chăm sóc đặc biệt.

Câu 3: (2 điểm)

Cho (X,Y) là vector ngẫu nhiên có hàm phân phối đồng thời

$$F(x,y)=\begin{cases}1-4^{-2x}-4^{-2y}+4^{-2x-2y} & khi \quad x\ge 0; y\ge 0\\ 0 & \text{ở những nơi khác}\end{cases}$$

a/ Hãy tìm hàm mật độ đồng thời $f(x,y)$.

b/ Hỏi X,Y có độc lập nhau hay không? Vì sao?

Câu 4: (2 điểm)

Một mẫu gồm 25 cá thể. Khi mang ra cân ta thu được các kết quả: 0,64 0,65 0,72 0,60 0,77 0,68 0,72 0,82 0,64 0,66 0,74 0,45 0,65 0,52 0,63 0,50 0,64 0,67 0,68 0,76 0,66 0,82 0,84 0,68 0,73 (kg).

a/ Hãy tìm KTC dành cho cân nặng trung bình của các cá thể, với ĐTC 95%.

b/ Để có độ chính xác là 0,02 thì cần có cỡ mẫu là bao nhiêu?

Câu 5: (2 điểm)

Khi sản xuất ra 400 sản phẩm bằng phương pháp thứ nhất thì ta thu được kết quả là 120 sản phẩm loại A, 70 sản phẩm loại C, còn lại là sản phẩm loại B. Còn khi sản xuất ra 500 sản phẩm bằng phương pháp thứ hai thì ta nhận được kết quả là 250 sản phẩm loại B, 170 sản phẩm loại A, còn lại là sản phẩm loại C. Như vậy, có thể kết luận được rằng hai phương pháp sản xuất là như nhau hay không, ứng với mức ý nghĩa $\alpha = 0{,}01$.

HẾT