BỘ CÔNG THƯƠNG

TRƯỜNG CAO ĐẲNG CÔNG NGHIỆP HUẾ

BÀI GIẢNG XÁC SUẤT THỐNG KÊ

Th.S. NGUYỄN HOÀNG ANH KHOA

Huế, tháng 08 năm 2014

CHƯƠNG 1. XÁC SUẤT

1.1. Giải tích tổ hợp

1.1.1. Quy tắc đếm

a) Quy tắc nhân:

Công việc có k giai đoạn. Giai đoạn i có n_i cách thực hiện thì có tất cả $n_1. n_2... n_k$ cách hoàn thành công việc

b) Quy tắc cộng:

Công việc được hoàn thành bởi 1 trong k hành động. Hành động i có n_i cách thực hiện thì có tất cả $n_1+ n_2+...+ n_k$ cách hoàn thành công việc

1.1.2. Chỉnh hợp, tổ hợp

a) Chỉnh hợp: Chỉnh hợp chập k của n phần tử là một bộ gồm k phần tử có thứ tự lấy từ n phần tử khác nhau ($1\le k\le n$).

Số chỉnh hợp chập k của n phần tử: $A_n^k = \dfrac{n!}{(n-k)!}$

b) Hoán vị của n phần tử: Hoán vị của n phần tử là một bộ sắp thứ tự của n phần tử khác nhau

Số hoán vị của n phần tử: $P_n = n!$

c) Tổ hợp: Tổ hợp chập k của n phần tử ($1\le k\le n$) là một bộ gồm k phần tử khác nhau lấy từ n phần tử khác nhau không kể thứ tự.

Số tổ hơp chập k của n phần tử: $C_n^k = \dfrac{n!}{(n-k)!k!}$

1.1.3. Nhị thức Newton:

$$\left(a+b\right)^n = \sum_{k=0}^{n} C_n^k a^{n-k} b^k$$

1.1.4. Các ví dụ

1. Có bao nhiêu cách xếp 12 sinh viên vào 4 lớp A, B, C, D sao cho mỗi lớp có 3 sinh viên.

2. Một chồng sách gồm có 3 cuốn sách Toán, 4 cuốn sách Lý và 5 cuốn sách Hóa khác nhau.

a) Có bao nhiêu cách xếp 12 cuốn sách đó theo từng môn.

b) Có bao nhiêu cách xếp 12 cuốn sách đó sao cho 4 sách Lý đặt kề nhau.

3. Có bao nhiêu cách phát 10 món quà khác nhau cho 3 người sao cho người nào cũng có ít nhất một món quà.

1.2. Phép thử - biến cố

1.2.1. Phép thử: Là hành động, thí nghiệm ... để nghiên cứu hiện tượng nào đó.

1.2.2. Biến cố: Là hiện tượng có thể xảy ra hay không xảy ra trong kết cục của một phép thử

Quy ước: Dùng chữ cái in hoa để kí hiệu cho biến cố

Ví dụ: Phép thử là gieo 1 con xúc xắc. Biến cố "xuất hiện mặt 3 chấm", "xuất hiện mặt có số chấm là số chẵn". . .

1.2.3. Các phép toán về biến cố

- **Biến cố chắc chắn Ω** : biến cố nhất định xảy ra khi thực hiện phép thử.
- **Biến cố không thể** $\varnothing$: biến cố không thể xảy ra khi thực hiện phép thử.
- **Biến cố tích** AB: biến cố xảy ra nếu A và B đồng thời xảy ra.
- **Biến cố tổng** A + B: biến cố xảy ra nếu ít nhất 1 trong 2 biến cố A,B xảy ra.
- **Quan hệ kéo theo** $A\subset B$: Nếu A xảy ra thì B xảy ra.
- **Biến cố đối lập:** biến cố đối lập của biến cố A là biến cố $\overline{A}$="A không xảy ra"
- **Biến cố xung khắc:** A và B gọi là xung khắc nếu $A.B=\varnothing$

1.3. Xác suất của biến cố

1.3.1. Định nghĩa xác suất cổ điển

Định nghĩa: Nếu trong một phép thử có n biến cố đồng khả năng, trong đó có m biến cố thuận lợi cho biến cố A thì tỉ số m/n gọi là xác suất của biến cố A, kí hiệu P(A). Vậy

$$P(A)=\frac{m}{n}$$

Trong đó m: số biến cố sơ cấp thuận lợi cho A, kí hiệu n(A)

n : số biến cố sơ cấp đồng khả năng, kí hiệu n(Ω)

$$P(A)=\frac{n(A)}{n(\Omega)}$$

Ví dụ 1: Gieo đồng thời hai xúc xắc cân đối và đồng chất. Tìm xác suất để tổng số chấm xuất hiện trên hai xúc xắc bằng 6.

Giải

Gọi A là biến cố tổng số chấm xuất hiện trên hai xúc xắc bằng 6

Số biến cố đồng khả năng n(Ω) = 6.6 = 36

Số biến cố thuận lợi cho A là n(A) = 5

Vậy $$P(A)=\frac{n(A)}{n(\Omega)}=\frac{5}{36}$$

1.3.2. Định nghĩa xác suất theo quan niệm thống kê

Thực hiện n lần một phép thử thấy có m lần xuất hiện biến cố A. Khi đó, tỉ số $f_n(A):=m/n$ gọi là tần suất xuất hiện biến cố A khi thực hiện n lần phép thử.

Nếu giới hạn $\lim\limits_{n\to\infty} f_n(A)$ tồn tại thì xác suất của biến cố A kí hiệu P(A) xác định bởi công thức: $P(A) = \lim\limits_{n\to\infty} f_n(A)$

Trong thực tế, khi n đủ lớn ta có: $P(A) \simeq f_n(A)$

1.3.3. Tính chất của xác suất

Cho A, B là các biến cố bất kỳ trong một phép thử ta có:

1. $0 \le P(A) \le 1$; $P(\varnothing) = 0$ và $P(\Omega) = 1$
2. Nếu $A.B = \varnothing$ thì $P(A + B) = P(A) + P(B)$
3. $P(\bar{A}) = 1 - P(A)$

1.4. Xác suất có điều kiện

1.4.1. Định nghĩa: Cho A, B là hai biến cố bất kỳ trong một phép thử và P(A)>0. Xác suất có điều kiện của biến cố A với điều kiện biến cố B đã xảy ra là một số ký hiệu là P(A/B) được xác định bởi công thức:

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

1.4.2. Biến cố độc lập:

Hai biến cố A và B gọi là độc lập nếu P(A/B) = P(A) và P(B/A) = P(B)

Các biến cố $A_1,A_2,..,A_n$ gọi là độc lập nếu A_i và A_j độc lập với mọi $i \neq j$.

1.5. Công thức tính xác suất

1.5.1. Công thức nhân:

Cho A, B là hai biến cố của một phép thử, ta có

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

Mở rộng:

$$P(A_1A_2A_3\ldots A_{n-1}A_n)=P(A_1)P(A_2/A_1)P(A_3/A_1A_2)\ldots P(A_n/A_1A_2A_3\ldots A_{n-1})$$

Đặc biệt, nếu $A_1, A_2,.., A_n$ độc lập từng đôi thì $P(A_1A_2..A_n) = P(A_1)P(A_2)..P(A_n)$

Ví dụ 2: Hai hộp chứa các quả cầu. Hộp thứ nhất chứa 3 quả đỏ và 2 quả xanh, hộp thứ hai chứa 4 quả đỏ và 6 quả xanh. Lấy ngẫu nhiên từ mỗi hộp một quả. Tính xác suất:

a) cả 2 quả đều đỏ.

b) cả 2 quả đều xanh.

c) hai quả khác màu

d) quả lấy từ hộp thứ nhất là quả màu đỏ, biết rằng 2 quả khác màu.

Giải

a) Gọi A là biến cố cả 2 quả đều đỏ

A_1 là biến cố quả lấy từ hộp 1 là quả màu đỏ

A_2 là biến cố quả lấy từ hộp 2 là quả màu đỏ

Ta có A_1, A_2 độc lập

$$P(A) = P(A_1A_2) = P(A_1)P(A_2) = \frac{3}{5}.\frac{4}{10} = \frac{6}{25} = 0,24$$

b)Gọi B là biến cố cả 2 quả đều xanh.

$$P(B) = P(\overline{A_1}.\overline{A_2}) = P(\overline{A_1}).P(\overline{A_2}) = \frac{2}{5}.\frac{6}{10} = \frac{6}{25} = 0,24$$

c) Gọi C là biến cố hai quả khác màu.

$$P(C) = P(\overline{A+B}) = 1 - P(A+B) = 1 - [P(A) + P(B)] = 0,52$$

d) Gọi D là biến cố quả lấy từ hộp thứ nhất là quả màu đỏ, biết rằng 2 quả khác màu.

$$P(D) = P(A_1/C) = \frac{P(A_1C)}{P(C)} = \frac{P(A_1\overline{A_2})}{P(C)} = \frac{\frac{3}{5}.\frac{6}{10}}{0,52} = \frac{9}{13}$$

1.5.2. Công thức cộng:

Cho A, B là hai biến cố của một phép thử, ta có

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

Mở rộng:

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\le i<j\le n} P(A_iA_j) + \sum_{1\le i<j<k\le n} P(A_iA_jA_k) + .. + (-1)^{n+1}P(A_1A_2..A_n)$$

Đặc biệt, nếu $A_iA_j = \varnothing$ với mọi $i \neq j$ thì $P(A_1+A_2+..+A_n)=P(A_1)+P(A_2)+..+P(A_n)$

Ví dụ 3: Phát ngẫu nhiên 9 món quà cho 3 người. Tính xác suất có ít nhất một người không nhận được quà.

1.5.3. Công thức xác suất đầy đủ, công thức bayes

Nhóm đầy đủ: $\{A_i \mid i = 1, 2, 3, ... n\}$ là một nhóm đầy đủ nếu

Giả sử $\{A_i \mid i = 1, 2, 3, ... n\}$ là một nhóm đầy đủ và A là một biến cố xảy ra chỉ khi một trong các biến cố A_i xảy ra, khi đó:

a. Công thức xác suất đầy đủ

$$P(A) = P(A_1)P(A/A_1) + P(A_2)P(A/A_2) + ... + P(A_n)P(A/A_n)$$

b. Công thức Bayes

$$P(A_i/A) = \frac{P(A_i)P(A/A_i)}{P(A)} = \frac{P(A_i)P(A/A_i)}{\sum_{k=1}^{n} P(A_k)P(A/A_k)}$$

Ví dụ 4: Một phân xưởng có số lượng nam công nhân gấp 4 lần số lượng nữ công nhân. Tỷ lệ công nhân tốt nghiệp THPT đối với nữ là 15%, nam là 25%. Chọn ngẫu nhiên 1 công nhân của phân xưởng này. Tính xác suất:

a) chọn được:

- nam công nhân
- nữ công nhân

b) chọn được công nhân đã tốt nghiệp THPT.

c) chọn được nam công nhân tốt nghiệp THPT.

d) chọn được công nhân nữ, biết rằng người này đã tốt nghiệp THPT.

Giải

a) Gọi A là biến cố chọn được công nhân nam

=> là $\overline{A}$ biến cố chọn được công nhân nữ

$$P(A) = \frac{4}{5} \text{ và } P(\overline{A}) = \frac{1}{5}$$

b) Gọi B là biến cố chọn được công nhân đã tốt nghiệp THPT.

Ta có A, $\overline{A}$ là nhóm đầy đủ nên

$$P(B) = P(A).P(B/A) + P(\overline{A}).P(B/\overline{A}) = \frac{4}{5}.0,25 + \frac{1}{5}.0,15 = 0,23$$

c) Gọi C là biến cố chọn được công nhân nam tốt nghiệp THPT.

$$P(C) = P(AB) = P(A).P(B/A) = \frac{4}{5}.0,25 = 0,2$$

d) Gọi D là biến cố chọn được công nhân nữ, biết rằng người này đã tốt nghiệp THPT.

$$P(D) = P(\overline{A}/B) = \frac{P(\overline{A}.B)}{P(B)} = \frac{P(\overline{A}).P(B/\overline{A})}{P(B)} = \frac{\frac{1}{5}.0,15}{0,23} = \frac{3}{23}$$

BÀI TẬP CHƯƠNG 1

Câu1: Hai bạn Đào và Mai học xa nhà. Xác suất để Đào và Mai về thăm nhà vào ngày chủ nhật tương ứng là 0,2 và 0,25. Tính xác suất vào ngày chủ nhật:

a) cả hai về thăm nhà.

b) cả hai không về thăm nhà.

c) có đúng 1 người về thăm nhà.

d) Mai về thăm nhà, biết có đúng một người về thăm nhà.

Câu 2: Một tín hiệu S được truyền từ điểm A đến điểm B. Tín hiệu sẽ được nhận tại B nếu cả hai công tắc I và II đều đóng. Giả sử rằng khả năng để công tắc thứ nhất và thứ hai đóng, tương ứng là 0,8 và 0,6. Cho biết hai công tắc hoạt động độc lập nhau. Tính xác suất:

a) tín hiệu được nhận tại B.

b) công tắc thứ I mở, biết rằng tại B không nhận được tín hiệu S.

c) công tắc thứ II mở, biết rằng tại B không nhận được tín hiệu S.

d) cả hai công tắc I và II mở, biết rằng tại B không nhận được tín hiệu S.

Câu 3: Có 3 hộp: mỗi hộp đựng 5 viên bi, trong đó hộp thứ i có i viên bi trắng (i = 1,2,3). Lấy ngẫu nhiên từ mỗi hộp ra một viên bi.

a) Tìm xác suất lấy được 3 viên bi trắng.

b) Tính xác suất lấy được đúng 2 viên bi trắng

c) Tính xác suất lấy được đúng 1 viên bi trắng

d) Nếu trong 3 bi lấy ra có đúng 1 bi trắng, tìm xác suất viên bi trắng đó là của hộp thứ nhất?

Câu 4: Ba người chơi bóng rổ, mỗi người ném một quả. Xác suất ném trúng rổ của mỗi người lần lượt là 0,5; 0,6 và 0,7. Tính xác suất:

a) cả 3 người đều ném trúng rổ.

b) có ít nhất một người ném trúng rổ.

c) có đúng một người ném không trúng rổ.

d) người thứ nhất ném trúng rổ, biết có đúng một người ném không trúng rổ.

Câu 5: Hai bạn Bình và Yên cùng dự thi môn xác suất thống kê một cách độc lập. Khả năng để Yên thi đạt môn này là 0,6 và xác suất để có ít nhất một trong hai bạn thi đạt là 0,9. Tính xác suất:

a) bạn Bình thi đạt.

b) cả hai bạn đều thi đạt.

c) có ít nhất một bạn thi hỏng.

Câu 6: Có hai chuồng gà: chuồng I có 12 con gà mái và 8 con gà trống; chuồng II có 15 con gà mái và 10 con gà trống. Quan sát thấy có 2 con gà chạy từ chuồng I sang chuồng II; sau đó, có 1 con gà chạy từ chuồng II ra ngoài. Tính xác suất:

a) hai con gà chạy từ chuồng I sang chuồng II là 2 con gà mái.

b) trong hai con gà chạy từ chuồng I sang chuồng II có 1 con gà trống và 1 con gà mái.

c) hai con gà chạy từ chuồng I sang chuồng II là 2 con gà trống.

d) con gà chạy từ chuồng II ra ngoài là con gà trống.

Câu 7: Có hai chuồng thỏ, chuồng I có 8 thỏ đen và 12 thỏ trắng; chuồng II có 6 thỏ đen và 15 thỏ trắng. Quan sát thấy từ chuồng I có 1 con thỏ chạy sang chuồng II; sau đó, từ chuồng II có 2 con thỏ chạy ra ngoài. Tính xác suất:

a) con thỏ từ chuồng I chạy sang chuồng II:

- là thỏ trắng.
- là thỏ đen.

b) hai con thỏ chạy từ chuồng II ra ngoài là hai con thỏ trắng.

c) trong 2 con thỏ chạy ra từ chuồng 2 có 1 con thỏ trắng và 1 thỏ đen.

d) hai con thỏ chạy từ chuồng II ra ngoài là hai con thỏ đen.

Câu 8: Hai xạ thủ cùng bắn vào một mục tiêu (một xạ thủ bắn một viên đạn). Biết xác suất bắn trúng mục tiêu của xạ thủ I và II lần lượt là 0,8 và 0,9.

a) Tính xác suất cả hai xạ thủ đều bắn trúng mục tiêu.

b) Tính xác suất có đúng một xạ thủ bắn trúng mục tiêu.

c) Biết có đúng một xạ thủ bắn trúng mục tiêu. Tính xác suất xạ thủ bắn I trúng mục tiêu.

d) Biết có đúng một xạ thủ bắn trúng mục tiêu, xạ thủ bắn trượt lần thứ nhất tiếp tục bắn lần thứ hai. Tính xác suất lần hai xạ thủ bắn trúng mục tiêu.

CHƯƠNG 2. BIẾN NGẪU NHIÊN

2.1. Khái niệm

2.1.1. Định nghĩa: Hàm số X xác định trên không gian biến cố sơ cấp Ω được gọi là biến ngẫu nhiên (BNN)

Ví dụ 1: Gọi X là số lần xuất hiện mặt sấp khi gieo 10 lần một đồng xu, khi đó X là một BNN và X có thể nhận các giá trị là 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Ví dụ 2: Gọi X là số hạt giống nảy mầm khi gieo n hạt, khi đó X là một BNN và X có thể nhận các giá trị là 0, 1, 2, 3, ..., n. Kí hiệu $X(\Omega) = \{1,1,2,\ldots,n\}$.

Ví dụ 3: Gọi X là thời gian sử dụng của bóng đèn (đơn vị giờ). Khi đó, X là BNN có thể nhận các giá trị trong khoảng $[0,+\infty)$

2.1.2. Phân loại biến ngẫu nhiên

Dựa vào tập giá trị của BNN người ta chia BNN thành hai loại là BNN rời rạc và BNN liên tục

Định nghĩa: BNN mà tập hợp các giá trị nó có thể nhận là một tập hữu hạn hoặc vô hạn đếm được gọi là biến ngẫu nhiên rời rạc, ngược lại gọi là BNN liên tục.

Ví dụ: Trong các ví dụ trên: BNN X trong ví dụ 1 và ví dụ 2 là BNN rời rạc, BNN X trong ví dụ 3 là BNN liên tục

2.2. Biến ngẫu nhiên rời rạc

2.2.1. Bảng phân phối xác suất: là bảng cho biết thông tin các giá trị có thể nhận và xác suất để nhận các giá trị đó.

Giả sử biến ngẫu nhiên X nhận các giá trị $x_1, x_2, x_3, \ldots x_n$ với các xác suất tương ứng là $P(X = x_i) = p_i$. Ta có bảng phân phối xác suất:

X	x_1	x_2	X_3	...	x_n
P	p_1	p_2	P_3	...	p_n

Chú ý: $p_1 + p_2 + \ldots + p_n = 1$.

Ví dụ 1: Một sinh viên làm 2 thí nghiệm A, B với xác suất thành công của các thí nghiệm tương ứng là 0,6 và 0,7. Gọi X là số lần sinh viên làm thí nghiệm thành công. Lập bảng phân phối xác suất của X.

<u>Giải</u>

Các gí trị X có thể nhận $X(\Omega) = \{0;1;2\}$

Gọi A là biến cố sinh viên làm thí nghiệm A thành công

B là biến cố sinh viên làm thí nghiệm B thành công

Ta có A, B độc lập

$P(X = 0) = P(\overline{A}.\overline{B}) = P(\overline{A}).P(\overline{B}) = 0,4.0,3 = 0,12$

$P(X = 1) = P(A.\overline{B} + \overline{A}.B) = P(A).P(\overline{B}) + P(\overline{A}).P(B) = 0{,}6.0{,}3 + 0{,}4.0{,}7 = 0{,}46$

$P(X = 2) = P(A.B) = P(A)P(B) = 0{,}6.0{,}7 = 0{,}42$

Bảng phân phối xác suất

X	0	1	2
P	0,42	0,46	0,42

2.2.2. Hàm phân phối

Hàm phân phối của biến ngẫu nhiên X kí kiệu F(x) được xác định bởi công thức

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

Tính chất: Giả sử F(x) là hàm phân phối của BNN X, ta có:

- F(x) là hàm không giảm trên R.
- $0 \le F(x) \le 1$, với mọi $x \in (-\infty; +\infty)$;
- $F(-\infty) = 0$; $F(+\infty) = 1$
- $P(a \le X < b) = F(b) – F(a)$

2.2.3. Các đặc trưng

1. **Mod:** Mốt của X kí hiệu ModX là giá trị x_i sao cho $P(X=x_i) = p_i$ lớn nhất

2. **Trung vị:** Trung vị của X kí hiệu MedX là giá trị x_i sao cho $F(x_i) \le 0{,}5$ và $F(x_{i+1}) > 0{,}5$.

3. **Kì vọng:** $EX = \sum_{i=1}^{n} x_i p_i$

Tính chất:

1. EC = C
2. EkX = kEX
3. $E(X \pm Y) = EX \pm EY$, trong đó X, Y là hai biến ngẫu nhiên độc lập.

4. **Phương sai:** $VX = \sum_{i=1}^{n} (x_i - EX)^2 p_i$

Tính chất:

1. VC = 0
2. $VkX = k^2VX$
3. $V(X \pm Y) = VX + VY$, trong đó X, Y là hai biến ngẫu nhiên độc lập.

Chú ý: $\boxed{VX = E(X^2) - (EX)^2}$ với $EX^2 = \sum_{i=1}^{n} x_i^2 p_i$

5. **Độ lệch chuẩn:** $\sigma(X) = \sqrt{VX}$

Ví dụ 2: Cho X là biến ngẫu nhiên có bảng phân phối xác suất:

X	-2	0	1	2	3
P	0,1	0,2	0,1	0,5	0,1

a) Tìm hàm phân phối xác suất của X .

b) Tính xác suất P ($0 \leq X < 3$).

c) Tính mốt, trung vị, kì vọng, phương sai của X.

Giải

a) Hàm phân phối xác suất

$$F(x) = \begin{cases} 0 & , \quad x \leq -2 \\ 0,1 & , \quad -2 < x \leq 0 \\ 0,3 & , \quad 0 < x \leq 1 \\ 0,4 & , \quad 1 < x \leq 2 \\ 0,9 & , \quad 2 < x \leq 3 \\ 1 & , \quad x > 3 \end{cases}$$

b) $P(0 \leq X < 3) = F(3) - F(0) = 0,9 - 0,1 = 0,8$

c) ModX = 2

MedX = 2

EX = -2.0,1 + 0.0,2 + 1.0,1 + 2.0,5 + 3.0,1 = 1,2

$EX^2 = (-2)^2.0,1 + 0^2.0,2 + 1^2.0,1 + 2^2.0,5 + 3^2.0,1 = 3,4$

$VX = EX^2 - (EX)^2 = 3,4 - 1,2^2 = 1,96$

2.3. Biến ngẫu nhiên liên tục

2.3.1. Hàm mật độ: Hàm f(x) được gọi là hàm mật độ của một BNN liên tục X nào đó, nếu nó thỏa mãn ba điều kiện sau:

i) $f(x) \geq 0$

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

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

2.3.2. Hàm phân phối xác suất

Hàm phân phối của biến ngẫu nhiên X kí kiệu F(x) được xác định bởi công thức

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

Nhận xét:

- F(x) là hàm liên tục

- Nếu X là BNN liên tục thì $P(X = x_0) = 0$.
- Nếu X là BNN thì $P(a\le X<b)=P(a<X\le b)=P(a<X<b)=P(a\le X\le b)=F(b)-F(a)$
- $f(x) = F'(x)$ tại những điểm f(x) liên tục

2.3.3. Các đặc trưng

1. Mod: Mốt của X, kí hiệu modX là giá trị làm hàm mật độ đạt giá trị lớn nhất

2. Trung vị: Trung vị của X, kí hiệu ModX là giá trị x* sao cho $F(x^*) = 0,5$.

3. Kì vọng: $E(X) = \int_{-\infty}^{+\infty} xf(x)dx$

Tính chất:

1. EC = C
2. EkX = kEX
3. $E(X \pm Y) = EX \pm EY$, trong đó X, Y là hai biến ngẫu nhiên độc lập.

4. Phương sai: $VX = \int_{-\infty}^{+\infty} (x - EX)^2 f(x)dx$

Chú ý: $\boxed{VX = E(X^2) - (EX)^2}$ trong đó $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx$

Tính chất:

1. VC = 0
2. $VkX = k^2VX$
3. $V(X \pm Y) = VX + VY$, trong đó X, Y là hai biến ngẫu nhiên độc lập.

Ví dụ 3: Cho biến ngẫu nhiên liên tục X có hàm phân phối xác suất như sau:

$$F(x) = \begin{cases} 0 & ;x < 0 \\ ax^3 + b & ;0 \le x \le 1 \\ 1 & ;1 < x \end{cases}$$

a. Xác định hàm mật độ f(x) của biến ngẫu nhiên X, biết f liên tục trên R\{0;1}.

b. Tính kì vọng, phương sai.

<u>Giải</u>

a. Vì F(x) là hàm liên tục nên $\begin{cases} \lim\limits_{x\to 0^+} F(x) = \lim\limits_{x\to 0^-} F(x) \\ \lim\limits_{x\to 1^+} F(x) = \lim\limits_{x\to 1^-} F(x) \end{cases} \Leftrightarrow \begin{cases} b = 0 \\ 1 = a + b \end{cases}$

$\Rightarrow$ a = 1 và b = 0

Hàm mật độ $f(x) = F'(x) = \begin{cases} 0 & ;x < 0 \vee x > 1 \\ 3x^2 & ;0 \le x \le 1 \end{cases}$

b. Tính EX; V(X).

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

$$= \int_{-\infty}^{0} xf(x)dx + \int_{0}^{1} xf(x)dx + \int_{1}^{+\infty} xf(x)dx$$

$$= \int_{0}^{1} xf(x)dx = \int_{0}^{1} x.3x^2dx = 3.\left.\frac{x^4}{4}\right|_0^1 = \frac{3}{4}$$

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

$$= \int_{0}^{1} x^2.3x^2dx$$

$$= \int_{0}^{1} 3x^4dx = 3.\left.\frac{x^5}{5}\right|_0^1 = \frac{3}{5}$$

$$VX = EX^2 - (EX)^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{80}$$

BÀI TẬP CHƯƠNG 2

Câu 1: Một kiện hàng có 10 sản phẩm, trong đó có 6 sản phẩm loại I và 4 sản phẩm loại II. Chọn ngẫu nhiên (đồng thời) từ kiện hàng ra 2 sản phẩm. Gọi X là số sản phẩm loại II được lấy ra. Lập bảng phân phối xác suất của X .

Câu 2: Kiện hàng I có 12 sản phẩm trong đó có 3 phế phẩm, kiện hàng II có 15 sản phẩm trong đó có 5 phế phẩm. Chọn ngẫu nhiên từ mỗi kiện hàng ra 1 sản phẩm. Gọi X là số sản phẩm tốt chọn được. Lập bảng phân phối xác suất của X .

Câu 3: Lô hàng I có 10 sản phẩm tốt và 2 phế phẩm, lô hàng II có 14 sản phẩm tốt và 5 phế phẩm. Chọn ngẫu nhiên từ lô hàng I ra 1 sản phẩm và bỏ vào lô hàng II. Sau đó, từ lô hàng II chọn ngẫu nhiên ra 1 sản phẩm. Gọi X là số sản phẩm tốt lấy ra từ lô hàng II. Lập bảng phân phối xác suất của X.

Câu 4: Cho X là biến ngẫu nhiên có bảng phân phối xác suất:

X	1	2	3
P	0,2	0,5	p

a) Xác định p.

b) Tính kì vọng và phương sai của biến ngẫu nhiên X.

Câu 5: Một sinh viên được làm thí nghiệm A tối đa 3 lần, nếu có 1 lần thành công thì dừng lại, xác suất thành công của mỗi lần thí nghiệm là 0,7. Gọi X là số lần sinh viên làm thí nghiệm. Lập bảng phân phối xác suất của X.

Câu 6: Một xạ thủ có 3 viên đạn, anh ta lần lượt bắn từng viên đạn vào một mục tiêu cho đến khi trúng mục tiêu hoặc hết đạn. Biết xác suất bắn trúng mục tiêu của mỗi viên đạn là 0,6. Gọi X là số viên đạn xạ thủ bắn. Lập bảng phân phối xác suất của biến ngẫu nhiên X.

Câu 7: Một hộp có 5 viên bi trong đó có 2 bi đỏ và 3 bi xanh. Lấy ngẫu nhiên 3 viên bi từ hộp bi trên, gọi X số bi đỏ được lấy ra. Lập bảng phân phối xác suất của X.

Câu 8: Một chùm có 5 chìa khóa trong đó có 3 chìa mở được ổ khóa. Một người mở ổ khóa bằng cách thử ngẫu nhiên từng chìa cho đến khi mở được ổ khóa (loại chìa đã thử ra khỏi chùm). Tính số lần thử trung bình để mở được ổ khóa.

Câu 9: Cho hàm số

$$f(x) = \begin{cases} 0 & , x \notin \left[-\pi/2; \pi/2\right] \\ a.\cos x & , x \in \left[-\pi/2; \pi/2\right] \end{cases}$$

a) Xác định a để f(x) là hàm mật độ của một BNN X nào đó.

b) Tìm hàm phân phối của X, tính $P(0 \le X \le \pi/4)$

c) Tính kì vọng và phương sai của biến ngẫu nhiên X

CHƯƠNG 3. CÁC PHÂN PHỐI XÁC SUẤT THƯỜNG DÙNG

3.1. Phân phối nhị thức

- **Dãy phép thử Becnulli:** Là dãy n phép thử độc lập thỏa mãn 2 điều kiện sau:

1. Mỗi phép thử chỉ xảy ra một trong hai biến cố A hoặc Ā;
2. P(A) = p không đổi trong mọi phép thử.

- **Định nghĩa:** Ký hiệu X là số lần xuất hiện biến cố A trong dãy n phép thử Becnulli với xác suất thành công trong mỗi phép thử là p. Khi đó ta nói X có phân phối nhị thức với tham số n, p. Kí hiệu X~B(n;p)

Ta có $P(X = k) = C_n^k p^k (1-p)^{n-k}$ với $k = 0, 1, ..., n$.

Ví dụ 1: Gieo liên tiếp ba lần đồng xu cân đối và đồng chất. Gọi X là số lần xuất hiện mặt sấp trong 3 lần gieo. Lập bảng phân phối xác suất của X. Tính xác suất để trong 3 lần gieo có nhiều nhất 1 lần xuất hiện mặt sấp.

<u>Giải</u>

Coi việc gieo 3 lần đồng xu như tiến hành 3 phép thử Bernoulli với xác suất xuất hiện mặt sấp là p = 1/2. Vậy X~B(3;1/2)

Bảng phân phối xác suất của X

X	0	1	2	3
P	0,125	0,375	0,375	0,125

Xác suất trong 3 lần gieo có nhiều nhất 1 lần xuất hiện mặt sấp là:

$$P(X \le 1) = P(X = 0) + P(X = 1) = 1/2$$

Đặc biệt: Khi n =1 (*hay X~B(1;p)*) ta nói X có phân phối ***"không - một"*** và kí hiệu X ~ A(p).

Bảng phân phối xác suất của BNN X ~ A(P).

X	0	1
P	1-p	p

Nếu X ~ A(p) thì EX = p và VX = p(1-p)

Định lý: Nếu X là biến ngẫu nhiên có phân phối nhị thức B(n, p) thì E(X) = np và V(X) = np(1-p)

<u>Chứng minh</u>

Gọi X_i là số lần xuất hiện biến cố A trong phép thử thứ i, i = 1,2 .., n.

Khi đó, $X_i \sim B(0;1)$ và $X = X_1 + X_2 + \ldots + X_n$

Vậy

$EX = E(X_1 + X_2 + \ldots + X_n) = np$

$VX = V(X_1 + X_2 + \ldots + X_n) = np(1\text{-}p)$

3.2. Phân phối Poisson.

Định nghĩa: Biến ngẫu nhiên X được gọi là có phân phối Poisson với tham số $\lambda > 0$, kí hiệu $X \sim P(\lambda)$ nếu tập giá trị của nó $X(\Omega)=\{0;1;2;\ldots;n;\ldots\}$ và:

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

Định lý: Nếu X là biến ngẫu nhiên có phân phối Poison tham số λ thì $EX=VX=\lambda$

Bài toán dẫn đến phân phối Poisson

Gọi X là số lần xuất hiện biến cố A trong khoảng thời gian $(t_1; t_2)$ thỏa 2 điều kiện:

- Số lần xuất hiện của một biến cố A trong khoảng thời gian $(t_1;t_2)$ không ảnh hưởng tới xác suất suất hiện biến cố A trong khoảng thời gian kế tiếp
- Số lần xuất hiện của biến cố A trong khoảng thời gian tỉ lệ thuận tỉ lệ thuận với độ dài của khoảng đó.

Khi đó $X \sim P(\lambda)$ với $\lambda=c(t_2-t_1)$, c là cường độ xuất hiện A (*số lần xuất hiện biến cố A trên một đơn vị thời gian*).

Ví dụ 2: Số xe máy cần qua trạm trung chuyển ở hầm Hải Vân là một biến ngẫu nhiên trung bình cứ 2 phút có 3 xe. Năng lực phục vụ của xe trung chuyển là 10 phút phục vụ được 20 xe. Tính xác suất có xe máy phải đợi hơn 10 phút mới được phụ vụ.

<u>*Giải*</u>

Số xe đến hầm trong 1 phút c = 3/2 =1,5. λ = 1,5.10 = 15

Gọi X là số xe đến hầm trong 10 phút, ta có $X \sim P(15)$

Xác suất có xe phải đợi hơn 10 phút là $P(X>10)=1 - P(X\le 10) =$

3.3. Phân phối Chuẩn

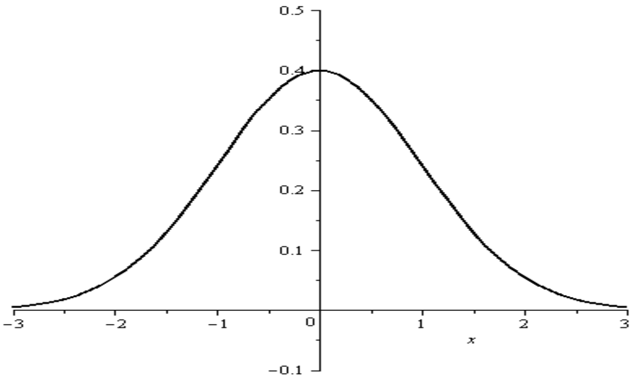
3.3.1. Định nghĩa: Biến ngẫu nhiên X gọi là có phân phối chuẩn, ký hiệu là $X \sim N(\mu,\sigma^2)$ nếu hàm mật độ của nó có dạng:

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

Đặc biệt, nếu $\mu = 0$ và $\sigma^2 = 1$ thì BNN X gọi là BNN có phân phối chuẩn tắc, ký hiệu $X \sim N(0;1)$. Khi đó hàm mật độ và hàm phân phối tương ứng có dạng:

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \quad \text{và} \quad F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

<u>Đồ thị:</u> hàm mật độ của biến ngẫu nhiên X~N(0;1)



3.3.2. Các định lí

Định lí: Nếu $X \sim N(\mu,\sigma^2)$ thì $EX = \mu$ và $VX = \sigma^2$

Định lí: Nếu $X_1 \sim N(\mu_1,\sigma_1^2)$ và $X_2 \sim N(\mu_2,\sigma_2^2)$ thì $X_1+X_2 \sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$ *(Xem [4] trang 100)*

3.3.3. Tính xác suất của phân phối chuẩn

TH1: $X \sim N(0;1)$

Ta có $P(a<X<b) = \varphi(b) - \varphi(a)$

với $\varphi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$ là hàm Laplace cho bởi bảng phụ lục 1.

Chú ý: $F(x) = \varphi(x) + 0,5$

$\varphi(-x) = -\varphi(x)$

TH2: $X \sim N(\mu,\sigma^2)$

Định lí: Nếu $X \sim N(\mu;\sigma^2)$ thì $aX+b \sim N(a\mu+b; a^2\sigma^2)$ *(Xem [1] trang 63)*

Hệ quả: $X \sim N(\mu,\sigma^2) \Leftrightarrow G = \frac{X-\mu}{\sigma} \sim N(0,1)$

Do đó, ta có: $$P(a < X < b) = \varphi\left(\frac{b-\mu}{\sigma}\right) - \varphi\left(\frac{a-\mu}{\sigma}\right)$$

Chú ý: ***Trong MS-Excel***: Nếu $X \sim N(0;1)$ thì F(**x)=Normsdist(x)**

3.3.4. Định lí giới hạn trung tâm

Định lí: Nếu X_i (i=1;2;..;n) là n BNN độc lập có cùng luật phân phối xác suất và $EX_i=\mu$, $VX_i=\sigma^2$ thì $X = \sum_{i=1}^{n} X_i \xrightarrow{F} N\left(n\mu; n\sigma^2\right)$ *(Xem [1] trang 66)*

Trong thực hành:

- Khi n đủ lớn ta xem X có phân phối $N(n\mu,n\sigma^2)$, viết $X \approx N(n\mu,n\sigma^2)$.
- Nếu $X \sim B(n;p)$ với $n>30$; $np>5$ và $n(1-p)>5$ ta xem $X \approx N(np;np(1-p))$

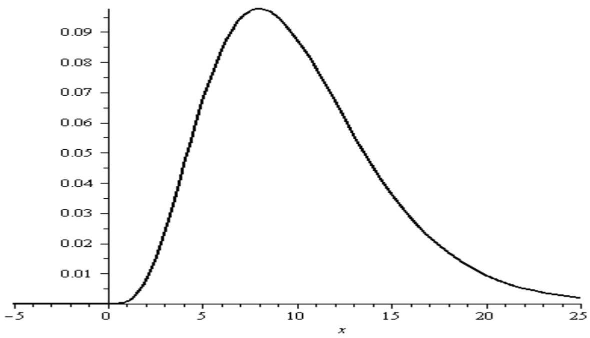
3.4. Phân phối khi bình phương

Định nghĩa: BNN X gọi là có phân phối "khi bình phương" với n bậc tự do, kí hiệu $X \sim \chi^2(n)$ nếu hàm mật độ xác suất của nó có dạng:

$$f(x) = \begin{cases} 0 & , \text{khi } x \le 0 \\ \dfrac{1}{2^{\frac{n}{2}}.\Gamma\left(\frac{n}{2}\right)} e^{\frac{x}{2}}.x^{\frac{n}{2}-1} & , \text{khi } x > 0 \end{cases}$$

trong đó $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$ là hàm Gamma

<u>Đồ thị</u>: hàm mật độ của BNN $X \sim \chi^2(10)$



Định lí: Nếu $X_1, X_2, ..., X_n$ là n BNN độc lập và có cùng phân phối chuẩn tắc thì

$$X = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$$

Định lí: Nếu $X \sim \chi^2(n)$ thì EX = n và VX = 2n

Định lí: Nếu $X \sim \chi^2(n)$ thì $X \xrightarrow{F} N(n;2n)$

Tính toán xác xuất cho phân phối **"Khi bình phương"** cho trong bảng phụ lục 2

Ví dụ 3: Cho $X \sim \chi^2(10)$. Tìm t biết P(X<t)=0,95.

Trong thực hành: Nếu $X \sim \chi^2(n)$ khi n đủ lớn (n>30) ta xem $X \approx N(n;2n)$

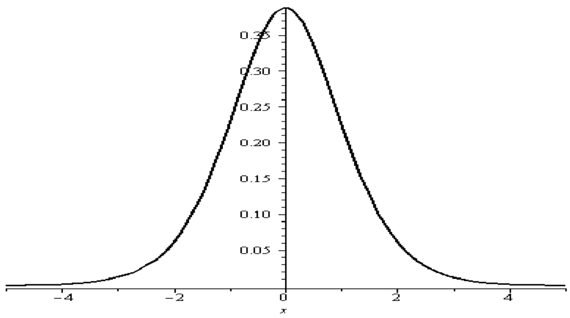
Chú ý: ***Trong MS-Excel*** ta có: $\boldsymbol{\chi_\alpha^2(k)}$**=chinv(α,k)**

3.5. Phân phối Student

Định nghĩa : BNN X gọi là có phân phối Student với n bậc tự do (kí hiệu X~t(n)) nếu hàm mật độ xác suất của nó có dạng :

$$f(x) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)}\left[1 + \frac{x^2}{n-1}\right]^{-\frac{n}{2}}$$

Đồ thị: hàm mật độ của BNN X~ t(10)



Định lí: Nếu X~N(0, 1) và Y~ χ^2(n) thì $Z = \dfrac{X}{\sqrt{\dfrac{Y}{n}}}$ ~t(n)

Tính toán xác xuất cho phân phối student cho trong bảng phụ lục 3

Ví dụ 4: Cho X ~ t(12). Tìm t biết P(X<t)= 0,99

Chú ý: ***Trong MS-Excel*** ta có: $\mathbf{t_\alpha(k) = Tinv(\alpha,k)}$

BÀI TẬP CHƯƠNG 3

Câu 1: Một xạ thủ bắn 10 viên đạn vào một mục tiêu, xác xuất trúng của mỗi viên là 0,6. Để phá hủy được mục tiêu cần phải bắn trúng ít nhất 3 viên, tính xác suất để mục tiêu bị phá hủy.

Câu 2: Trọng lượng sản phẩm X (đơn vị gam) do một máy tự động sản xuất ra có phân phối chuẩn X~N(100;2,56). Sản phẩm được coi là đạt kĩ thuật nếu trọng lượng của nó đạt từ 98 đến 102 gam.

a) Tìm tỉ lệ sản phẩm đạt tiêu chuẩn kĩ thuật của nhà máy

b) Cho máy sản suất 100 sản phẩm, tính xác suất có trên 90 sản phẩm đạt kĩ thuật.

Chương 4. LÝ THUYẾT MẪU

4.1. Đám đông, mẫu ngẫu nhiên

4.1.1. Đám đông, BNN của đám đông

Đám đông là tập hợp tất cả các đối tượng cần nghiên cứu

Dấu hiệu cần nghiên cứu thay đổi qua các phần tử của đám đông gọi là biến ngẫu nhiên của đám đông.

Ví dụ: Nghiên cứu về trọng lượng của các lon sữa A trên thị trường ở TP Huế

Đám đông là: Tập hợp các trên thị trường tại TP Huế

Biến ngẫu nhiên: trọng lượng của các lon sữa

4.1.2. Mẫu ngẫu nhiên, mẫu cụ thể, thống kê

Trong thực tế ta thường không thể nghiên của tất cả các phần tử của đám đông, để nghiên cứu các dấu hiệu của đám đông ta dùng phương pháp chọn mẫu.

Mẫu là một bộ phận của đám đông phản ảnh được các tính chất của đám đông

a) Mẫu ngẫu nhiên: là tập hợp n biến ngẫu nhiên $(X_1, X_2\ldots, X_n)$ độc lập có cùng luật phân phối với biến NN của đám đông

b) Cách chọn mẫu ngẫu nhiên: Chọn ngẫu nhiên đơn giản, chọn mẫu theo nhóm, chọn theo ý kiến chuyên gia …

Chú ý: Trong trong môn học này ta chỉ xét cách chọn mẫu sao cho mỗi phần tử của đám đông đều có khả năng được chọn là như nhau

c) Mẫu cụ thể: Giả sử X_i nhận giá trị là x_i với $i = 1,2,..,n$. Khi đó, $(x_1,x_2,..,x_n)$ gọi là một mẫu cụ thể kíc h thước n.

d) Thống kê: Một hàm của mẫu ngẫu nhiên $\varphi(X_1,X_2\ldots, X_n)$ gọi là một thống kê

4.1.3. Cách trình bày mẫu ngẫu nhiên cụ thể

Trường hợp 1: Mẫu nhỏ hoặc nhận ít giá trị ta dùng bảng phân phối tần số thực nghiệm dạng

X	a_1	a_2	…	a_k
Tần số	n_1	n_2	…	n_k

trong đó, a_i có n_i giá trị với $i = 1,2, \ldots,k$.

Chú ý: $n_1+n_2+\ldots+n_k=n$.

Đặt $f_i=n_i/n$ ta có bảng PP tần suất

X	a_1	a_2	…	a_k
Tần suất	f_1	f_2	…	f_k

Chú ý: $f_1+f_2+\ldots+f_n=1$

Trường hợp 2: Mẫu lớn và các giá trị phân tán ta dùng bảng phân phối ghép lớp dạng

X	b_0-b_1	b_1-b_2	…	b_{k-1}-b_k
Tần số	n_1	n_2	…	n_k

Chú ý: b_i thuộc khoảng b_{i-1}-b_i . Chọn số khoảng k sao cho $2^{k-1}<n<2^k$

Ngoài ra ta có thể dùng các loại biểu đồ

4.2. Các đặc trưng mẫu

Giả sử đám đông X có EX = μ, VX = σ^2 và p là tỉ lệ các phần tử có tính chất A của đám đông.

Cho $(X_1,\ldots, X_n)$ là mẫu ngẫu nhiên kích thước n và $(x_1,x_2,..,x_n)$ là một mẫu cụ thể kích thước n.

4.2.1. Tỉ lệ mẫu

a) Định nghĩa

Tỉ lệ mẫu $F=\frac{1}{n}\sum_{i=1}^{n}X_i$, tỉ lệ mẫu thực nghiệm $f=\frac{1}{n}\sum_{i=1}^{n}x_i$

b) Các số đặc trưng của tỉ lệ mẫu

- Kì vọng $EF = p$

- Phương sai $VF=\frac{p(1-p)}{n}$

c) Phân phối xác suất của tỉ lệ mẫu

- Khi n đủ lớn theo định lí giới hạn trung tâm ta có $F\simeq N\left(p,\frac{p(1-p)}{n}\right)$ hay

$$G=\frac{F-p}{\sqrt{p(1-p)}}\sqrt{n}\simeq N(0;1)$$

- Trong thực hành khi n đủ lớn và p chưa biết ta xem: $G=\frac{F-p}{\sqrt{f(1-f)}}\sqrt{n}\simeq N(0;1)$

4.2.2. Trung bình mẫu

a) Định nghĩa

Trung bình mẫu: $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

Trung bình mẫu thực nghiệm $\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{n}\sum_{i=1}^{k}a_in_i=\sum_{i=1}^{k}a_if_i$

Chú ý: Trường hợp dùng bảng phân phối tần số (suất) ghép lớp thì $a_i=(b_{i-1}+b_i)/2$

b) Các số đặc trưng

- Kì vọng $E\overline{X} = \mu$

- Phương sai $V\overline{X} = \sigma^2/n$

c) Phân phối xác suất của trung bình mẫu

- Trường hợp $X \sim N(\mu,\sigma^2)$ thì $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ hay $G = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \simeq N(0;1)$

- Trường hợp n > 30 theo định lí giới hạn trung tâm ta có:

$$G = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \simeq N(0;1) \text{ hay } G = \frac{\overline{X}-\mu}{s}\sqrt{n} \simeq N(0;1)$$

4.2.3. Phương sai mẫu

Phương sai mẫu : khi μ biết: $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \mu\right)^2$

khi μ chưa biết: $\overline{S}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

4.2.4. Phương sai mẫu hiệu chỉnh, độ lệch tiêu chuẩn

a) Định nghĩa

Phương sai mẫu hiệu chỉnh: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

Phương sai mẫu hiệu chỉnh thực nghiệm:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = \frac{1}{n-1}\sum_{i=1}^{k}n_i\left(a_i - \overline{x}\right)^2$$

Độ lệch tiêu chuẩn: $s = \sqrt{s^2}$

Chú ý: $s^2 = \frac{n}{n-1}\left[\overline{x^2} - \left(\overline{x}\right)^2\right]$ trong đó $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n}x_i^2$

b) Số đặc trưng: $ES^2 = \sigma^2$

c) Phân phối xác suất của phương sai mẫu hiệu chỉnh

- Nếu $X \sim N(\mu,\sigma^2)$ và μ đã biết thì $G = \frac{n.S^2}{\sigma^2} \sim \chi^2(n)$

- Nếu $X \sim N(\mu,\sigma^2)$ và μ chưa biết thì $G = \frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$

Hệ quả: Nếu $X \sim N(\mu,\sigma^2)$ và σ^2 chưa biết thì $G = \frac{\overline{X}-\mu}{S}\sqrt{n} \sim t(n-1)$

BÀI TẬP CHƯƠNG 4

Câu 1: Tỉ lệ phế phẩm của lô hàng là 5%. Kiểm tra ngẫu nhiên 500 sản phẩm của lô hàng nếu có từ 30 phế phẩm trở lên thì lô hàng không được phép xuất khẩu. Tính xác suất lô hàng được xuất khẩu.

Câu 2: Kiểm tra trọng lượng của một nhóm sinh viên nam được kết quả như sau

X (kg)	40-45	45-50	50-55	55-60	60-65	65-70
Tần số	5	15	20	15	3	2

Tính trung bình mẫu thực nghiệm, phương sai mẫu hiệu chỉnh thực nghiệm

Câu 3: Thời gian của một cuộc điện thoại đường dài tại một tổng đài là biến ngẫu nhiên có phân phối chuẩn trung bình 8 phút, độ lệch tiêu chuẩn 2 phút. Chọn ngẫu nhiên một mẫu 25 cuộc điện thoại đường dài ở tổng đài

a) Tìm độ lệch tiêu chuẩn của trung bình mẫu

b) Tính xác suất để trung bình mẫu từ 7,8 đến 8,2 phút.

Chương 5. ƯỚC LƯỢNG THAM SỐ

5.1. Khái niệm

Giả sử khi nghiên cứu ĐLNN X và biết được phân phối của X thuộc một loại phân phối nào đó (chẳng hạn biết X có phân phối chuẩn hoặc biết X có phân phối Poisson, ... nhưng lại không biết các tham số). Muốn xác định hoàn toàn phân phối của X ta phải xác định được các giá trị tham số của phân phối đó. Chính vì vậy, việc đi tìm ước lượng cho các tham số của phân phối là cần thiết.

5.2. Ước lượng điểm

5.2.1. Khái niệm: Cho mẫu ngẫu nhiên $(X_1, X_2, ..., X_n)$ của ĐLNN X, giả sử θ là tham ẩn cần ước lượng. Khi đó ước lượng điểm của tham số θ là ĐLNN $T_n = \varphi(X_1, X_2, ..., X_n)$ chỉ phụ thuộc vào $(X_1, X_2, ..., X_n)$.

5.2.2. Các tiêu chuẩn ước lượng

a) Ước lượng không chệch: Ước lượng T_n của tham số θ được gọi là ước lượng không chệch nếu $ET_n = \theta$

Ta có F, $\overline{X}$, S^2 là ước lượng không chệch cho p, μ, σ^2.

b) Ước lượng vững: Ước lượng T_n của tham số θ được gọi là ước lượng vững nếu thỏa mãn điều kiện

$$\lim_{n\to+\infty} P\left(\left|T_n - \theta\right| < \varepsilon\right) = 1, \quad \forall \varepsilon > 0.$$

Ta có F, $\overline{X}$, S^2 ước lượng vững cho p, μ, σ^2.

c) Ước lượng hiệu quả:

Định nghĩa: Thống kê T_n gọi là ước lượng hiệu quả của tham số θ nếu nó là ước lượng không chệch và có phương sai bé nhất trong các ước lượng không chệch của tham số θ.

Định lí: Nếu $ET_n = \theta$ và $VT_n = \dfrac{1}{nE\left(\dfrac{\partial \ln f(X,\theta)}{\partial\theta}\right)^2}$ thì T_n là ước lượng hiệu quả của tham số θ. *(xem [1] trang 129)*

Ta có F, $\overline{X}$, S^2 là ước lượng hiệu quả cho p, μ, σ^2.

5.3. Ước lượng khoảng:

Khái niệm

Khoảng (θ_1, θ_2) được gọi là khoảng ước lượng của θ với độ tin cậy $1 - \alpha$ nếu $P(\theta_1 < \theta < \theta_2) = 1 - \alpha$.

$\theta_2 - \theta_1 = 2\varepsilon$: ε gọi là độ chính xác của khoảng tin cậy

Bài toán: Từ mẫu $(x_1,x_2,..,x_n)$ tìm khoảng $(\theta_1; \theta_2)$ sao cho $P(\theta_1 < \theta < \theta_2) = 1 - \alpha$.

Các bước giải bài toán:

B1. Chọn thống kê $G(\theta,X_1, X_2, ..., X_n)$ có phân phối xác định

B2. Lấy t_1, t_2 sao cho $P(t_1<G<t_2)=1-\alpha$

B3. Biến đổi $t_1<G<t_2 \Leftrightarrow \theta_1(X_1, X_2, ..., X_n) < \theta < \theta_2(X_1, X_2, ..., X_n)$

Vận dụng nguyên lí xác suất lớn

Kết luận với độ tin cậy 1-α khoảng tin cậy là $(\theta_1(x_1,x_2,..,x_n); \theta_2(x_1,x_2,..,x_n))$

5.3.1. Ước lượng khoảng cho tỉ lệ

Khi n đủ lớn *(nf>10 và n(1-f)>10)* ta có

$$G = \frac{F-p}{\sqrt{f(1-f)}}\sqrt{n} \simeq N(0;1)$$

Khoảng tin cậy đối xứng

n = …. ; f = …………

Với độ tin cậy 1-α = … $\Rightarrow$ α = … $\Rightarrow t_{\alpha/2} = \varphi^{-1}(0,5-\alpha/2)$ → bảng phụ lục 1

Độ chính xác $\varepsilon = t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}}$

Khoảng tin cậy (f-ε; f+ε)

Ví dụ 1: Để ước lượng tỉ lệ người Việt Nam có nhóm máu O. Người ta kiểm tra ngẫu nhiên 1000 người được kết quả có 360 người có nhóm máu O.

a) Với độ tin cậy 95% hãy ước lượng tỉ lệ người Việt Nam có nhóm máu O.

b) Với tỉ lệ mẫu và độ tin cậy như câu a. Nếu muốn khoảng ước lượng có độ chính xác ε = 0,01 thì cần lấy mẫu gồm bao nhiêu người.

<u>Giải</u>

a) Ta có n = 1000 ; f = 0,36 ; 1 - α = 0,95 => α = 0,05

Mức phân vị $t_{\alpha/2} = \varphi^{-1}(0,5 - \alpha/2) = \varphi^{-1}(0,475) = 1,96$

Độ chính xác $\varepsilon = t_{\frac{\alpha}{2}}\frac{\sqrt{f(1-f)}}{\sqrt{n}} = 1,96\sqrt{\frac{0,36.(1-0,36)}{1000}} = 0,03$

Khoảng tin cậy (f - ε ; f + ε) = (0,33; 0,39)

b) Ta có $\varepsilon = t_{\frac{\alpha}{2}}\frac{\sqrt{f(1-f)}}{\sqrt{n}}$

=> $n = \left(t_{\frac{\alpha}{2}}\frac{\sqrt{f(1-f)}}{\varepsilon}\right)^2$

= 8851

Vậy cần chọn mẫu khoảng 8851 người

5.3.2. Ước lượng khoảng cho kỳ vọng

Trường hợp: $n > 30$

- Nếu σ^2 đã biết $G = \dfrac{\overline{X}-\mu}{\sigma}\sqrt{n} \simeq N(0;1)$

- Nếu σ^2 chưa biết $G = \dfrac{\overline{X}-\mu}{s}\sqrt{n} \simeq N(0;1)$

Trường hợp: $n \le 30$

- Nếu $X \sim N(\mu,\sigma^2)$ và σ^2 đã biết thì $G = \dfrac{\overline{X}-\mu}{\sigma}\sqrt{n} \sim N(0;1)$

- Nếu $X \sim N(\mu,\sigma^2)$ và σ^2 chưa biết thì $G = \dfrac{\overline{X}-\mu}{S}\sqrt{n} \sim t(n-1)$

Các bước ước lượng đối xứng cho kì vọng

	σ^2 biết	σ^2 **chưa** biết
n > 30	$n = \ldots$; $1 - \alpha = \ldots$ Mức phân vị: $t_{\alpha/2} = \ldots$ Độ chính xác: $\varepsilon = t_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}} = \ldots$ Khoảng tin cậy: $(\overline{x}-\varepsilon;\overline{x}+\varepsilon) = \ldots$	$n = \ldots$; $1 - \alpha = \ldots$ Mức phân vị: $t_{\alpha/2} = \ldots$ Độ chính xác: $\varepsilon = t_{\frac{\alpha}{2}}\dfrac{s}{\sqrt{n}} = \ldots$ Khoảng tin cậy: $(\overline{x}-\varepsilon;\overline{x}+\varepsilon) = \ldots$
n ≤ 30 X có phân phối chuẩn	$n = \ldots$; $1 - \alpha = \ldots$ Mức phân vị: $t_{\alpha/2} = \ldots$ Độ chính xác: $\varepsilon = t_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}} = \ldots$ Khoảng tin cậy: $(\overline{x}-\varepsilon;\overline{x}+\varepsilon) = \ldots$	$n = \ldots$; $1 - \alpha = \ldots$ Mức phân vị: $t_{\alpha/2}(n-1) = \ldots$ Độ chính xác: $\varepsilon = t_{\frac{\alpha}{2}}(n-1).\dfrac{s}{\sqrt{n}} = \ldots$ Khoảng tin cậy: $(\overline{x}-\varepsilon;\overline{x}+\varepsilon) = \ldots$

Ví dụ 2: Để ước lượng khối lượng trung bình của các bao gạo trong kho, người ta cân ngẫu nhiên 21 bao gạo trong kho được kết quả như sau: (đơn vị: kg)

50	48	49	48	50	51	50
49	50	51	50	52	51	51
50	51	50	49	51	50	49

a) Lập bảng phân phối tần số mẫu thực nghiệm, tính trung bình mẫu thực nghiệm, phương sai mẫu hiệu chỉnh thực nghiệm.

b) Với độ tin cậy 95% hãy tìm khoảng tin cậy cho khối lượng trung bình của các bao gạo trong kho. Biết khối lượng các bao gạo trong có phân phối chuẩn.

<u>*Giải*</u>

a) Bảng phân phối tần số mẫu thực nghiệm

X	48	49	50	51	52
Tần số	2	4	8	6	1

$$\overline{x} = \frac{1}{21}\left(48.2 + 49.4 + 50.8 + 51.6 + 52.1\right) = 50$$

$$\overline{x^2} = \frac{1}{21}\left(48^2.2 + 49^2.4 + 50^2.8 + 51^2.6 + 52^2.1\right) = \frac{52522}{21}$$

$$s^2 = \frac{21}{20}\left(\frac{52522}{21} - 50^2\right) = 1,1$$

b) n = 21 < 30

Độ tin cậy 1 - α = 0,95 => α = 0,05

Mức phân vị $t_{0,025}(20) = 2,0860$

Độ chính xác $\varepsilon = 2,0860.\sqrt{\frac{1,1}{21}} = 0,4774$

Khoảng tin cậy (49,5226; 50,4774)

5.3.3. Ước lượng khoảng cho phương sai

- Nếu $X \sim N(\mu,\sigma^2)$ và μ đã biết thì $G = \frac{nS^2}{\sigma^2} \sim \chi^2(n)$

- Nếu $X \sim N(\mu,\sigma^2)$ và μ chưa biết thì $G = \frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$

Các bước ước lượng cho phương sai

X có phân phối chuẩn

μ biết	μ **chưa** biết
n = … 1 - α = … => α = … Mức phân vị: $\chi^2_{\frac{\alpha}{2}}(n) = ...$ $\chi^2_{1-\frac{\alpha}{2}}(n) = ...$ Khoảng tin cậy $\left(\frac{ns^2}{\chi^2_{\frac{\alpha}{2}}(n)};\frac{ns^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right)$	n = … 1 - α = … => α = … Mức phân vị: $\chi^2_{\frac{\alpha}{2}}(n-1) = ...$ $\chi^2_{1-\frac{\alpha}{2}}(n-1) = ...$ Khoảng tin cậy $\left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)};\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$

Ví dụ 3: Tuổi thọ của một loại thiết bị A (đơn vị: tháng) có phân phối chuẩn, người ta điều tra ngẫu nhiên 15 thiết bị loại này được kết quả:

23	25	23	23	22
24	26	24	24	22
24	25	25	23	24

a) Tính trung bình mẫu thực nghiệm, phương sai mẫu hiệu chỉnh thực nghiệm.

b) Với độ tin cậy 98% hãy tìm khoảng tin cậy cho phương sai của tuổi thọ của loại thiết bị A.

<u>*Giải*</u>

a) Bảng phân phối tần số thực nghiệm

X	22	23	24	25	26
Tần số	2	4	5	3	1

$\overline{x} = 23{,}8$

$\overline{x^2} = 1703/3$

$s^2 = 46/35$

b) n = 15

Độ tin cậy 1 - α = 0,98 => α = 0,02

Mức phân vị $\chi^2_{0,01}(14) = 29{,}1412$

$\chi^2_{1-0,01}(14) = 4{,}6604$

Khoảng tin cậy: $\left(\dfrac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}; \dfrac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right) = \left(\dfrac{14.\frac{46}{35}}{29{,}1412}; \dfrac{14.\frac{46}{35}}{4{,}6604}\right) = (0{,}63;\ 3{,}95).$

BÀI TẬP CHƯƠNG 5

Câu 1: Kiểm tra ngẫu nhiên 500 sản phẩm của lô hàng thấy có 15 phế phẩm. Với độ tin cậy 98% hãy ước lượng tỉ lệ phế phẩm của lô hàng.

Câu 2: Tuổi thọ của một loại thiết bị A (đơn vị: tháng) là biến ngẫu nhiên có phân phối chuẩn, người ta điều tra ngẫu nhiên tuổi thọ của 21 thiết bị loại này được kết quả:

34	34	35	35	34	36	35
33	34	33	36	35	35	36
37	35	36	36	35	36	35

a) Lập bảng phân phối tần số thực nghiệm, tính trung bình mẫu thực nghiệm, tính phương sai mẫu hiệu chỉnh thực nghiệm.

b) Với độ tin cậy 95%, tìm khoảng tin cậy đối xứng cho tuổi thọ trung bình của loại thiết bị A.

c) Với độ tin cậy 95%, tìm khoảng tin cậy cho phương sai của tuổi thọ của loại thiết bị A.

Câu 3: Tuổi thọ của một loại thiết bị A (đơn vị: tháng) là biến ngẫu nhiên có phân phối chuẩn, người ta điều tra ngẫu nhiên tuổi thọ của 40 thiết bị loại này được kết quả:

40	41	41	39	41	42	38	38	39	39
39	41	41	41	39	42	41	40	40	40
41	40	39	39	40	40	41	40	40	39
40	40	41	39	40	40	41	40	40	38

a) Lập bảng phân phối tần số thực nghiệm, tính trung bình mẫu thực nghiệm, tính phương sai mẫu hiệu chỉnh thực nghiệm.

b) Với độ tin cậy 95%, tìm khoảng tin cậy đối xứng cho tuổi thọ trung bình của loại thiết bị A.

Câu 4: Một cửa hàng vật liệu xây dựng nhập xi măng từ cơ sở sản xuất A. Trọng lượng của các bao xi măng là biến ngẫu nhiên có phân phối chuẩn. Điều tra ngẫu nhiên trọng lượng X(kg) của 20 bao xi măng từ lô hàng mới nhập về, thu được số liệu sau:

X	49 - 49.5	49.5 - 50	50 – 50.5	50.5 - 51
Số bao	3	7	9	1

a) Tính trọng lượng trung bình và phương sai hiệu chỉnh của các bao xi măng trên.

b) Với độ tin cậy 95%, tìm khoảng ước lượng đối xứng của trọng lượng trung bình của các bao xi măng.

Chương 6. KIỂM ĐỊNH GIẢ THUYẾT THỐNG KÊ

6.1. Khái niệm

6.1.1. Giả thuyết thống kê:

Giả thuyết thống kê là giả thuyết nói về:

- Các tham số đặc trưng của biến ngẫu nhiên gốc của đám đông như tỉ lệ f, trung bình μ, phương sai σ^2.
- Dạng quy luật phân phối của biến ngẫu nhiên gốc của đám đông
- Tính độc lập của đám đông.

Trong chương trình của môn học này ta chỉ đề cập đến giả thuyết thống kê nói về các tham số đặc trưng của biến ngẫu nhiên gốc của đám đông.

6.1.2 Các bước cần thiết để kiểm định một giả thuyết thống kê

- Phát biểu giả thuyết H_0 và đối thuyết H_1
- Chọn thống kê G có phân phối xác định
- Thiết lập miền bác bỏ W_α (là miền sao cho $P\{G \in W_\alpha | H_0 \text{ đúng}\} = \alpha$)
- Kiểm tra $g = G(x_1;x_2;\ldots;x_n;\ldots) \in W_\alpha$ bác bỏ H_0 hoặc $g \notin W_\alpha$ chấp nhận H_0.

Khi ta chấp nhận hay bác bỏ H_0 ta mắc phải các sai lầm sau:

- **Sai lầm loại 1:** H_0 đúng mà ta bác bỏ. Xác suất biến cố này là $P\{G \in W_\alpha | H_0 \text{ đúng}\} = \alpha$
- **Sai lầm loại 2:** H_0 sai mà ta chấp nhận. Xác suất biến cố này là $P\{G \notin W_\alpha | H_0 \text{ sai}\} = \beta$

6.2. Kiểm định giả thuyết về tỉ lệ của đám đông

Bài toán: Giả sử p là tỉ lệ của đám đông X, chưa biết. Với mức ý nghĩa α, kiểm định giả thuyết: H_0: $p = p_0$ (p_0 đã biết)

Khi $np_0 \geq 5$ và $n(1-p_0) \geq 5$, ta có

$$G = \frac{F - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} \simeq N(0;1)$$

+ Khi H_1: $p \neq p_0$ miền bác bỏ $W_\alpha = \{|g| > t_{\alpha/2}\}$

+ Khi H_1: $p > p_0$ miền bác bỏ $W_\alpha = \{ g > t_\alpha \}$

+ Khi H_1: $p < p_0$ miền bác bỏ $W_\alpha = \{ g < - t_\alpha \}$

Ví dụ 1: Ở một nước, một đảng chính trị tuyên bố rằng 45% cử tri sẽ bỏ phiếu cho ông A là ứng cử viên của họ. Chọn ngẫu nhiên 200 người hỏi ý kiến có 80 người sẽ bầu cho ông A. Với mức ý nghĩa 5% hãy cho kết luận về tuyên bố trên.

Giải

Giả thuyết H_0: $p = 0,45$, đối thuyết H_1: $p \neq 0,45$

Mức ý nghĩa $\alpha = 0,05 \Rightarrow t_{\alpha/2} = 1,96$. Miền bác bỏ $W_\alpha = \{|g| > 1,96\}$

Ta có $g = \dfrac{f - P_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} = \dfrac{0,4-0,45}{\sqrt{0,45.(1-0,45)}}\sqrt{200} = -1,42 \notin W_\alpha$

Vậy chưa đủ cơ sở để bác bỏ tuyên bố trên.

6.3. Kiểm định giả thuyết về trung bình của đám đông

Bài toán: Giả sử μ là kì vọng (trung bình) của đám đông X, chưa biết. Với mức ý nghĩa α, kiểm định giả thuyết H_0: $\mu = \mu_0$ (μ_0 đã biết)

Khi H_0 đúng ta sử dụng các tiêu chuẩn kiểm định G trong các trường hợp sau:

Trường hợp 1: $n > 30$, σ^2 đã biết ta có $G = \dfrac{\overline{X} - \mu_0}{\sigma}\sqrt{n} \simeq N(0;1)$

+ Khi H_1: $\mu \neq \mu_0$ miền bác bỏ $W_\alpha = \{|g| > t_{\alpha/2}\}$

+ Khi H_1: $\mu > \mu_0$ miền bác bỏ $W_\alpha = \{ g > t_\alpha \}$

+ Khi H_1: $\mu < \mu_0$ miền bác bỏ $W_\alpha = \{ g < -t_\alpha \}$

Trường hợp 2: $n > 30$, σ^2 chưa biết: $G = \dfrac{\overline{X} - \mu_0}{s}\sqrt{n} \simeq N(0;1)$

Miền bác bỏ như trường hợp 1

Trường hợp 3: $n \leq 30$, $X \sim N(\mu,\sigma^2)$ và σ^2 đã biết thì $G = \dfrac{\overline{X} - \mu_0}{\sigma}\sqrt{n} \sim N(0;1)$

Miền bác bỏ như trường hợp 1

Trường hợp 4: $n \leq 30$, $X \sim N(\mu,\sigma^2)$ và σ^2 chưa biết thì $G = \dfrac{\overline{X} - \mu_0}{S}\sqrt{n} \sim t(n-1)$

+ Khi H_1: $\mu \neq \mu_0$ miền bác bỏ $W_\alpha = \{|g| > t_{\alpha/2}(n-1)\}$

+ Khi H_1: $\mu > \mu_0$ miền bác bỏ $W_\alpha = \{ g > t_\alpha(n-1) \}$

+ Khi H_1: $\mu < \mu_0$ miền bác bỏ $W_\alpha = \{ g < -t_\alpha(n-1) \}$

Ví dụ 2: Trọng lượng của các bao gạo là biến ngẫu nhiên có phân phối chuẩn với trọng lượng trung bình bình thường là 50kg. Nghi ngờ máy đóng bao gạo làm việc không bình thường làm cho trọng lượng của bao gạo có xu hướng giảm sút. Người ta cân thử 25 bao được khối lượng trung bình là 49,25kg và độ lệch chuẩn 0,81 kg. Với mức ý nghĩa 1% hãy cho kết luận về nghi ngờ trên.

Giải

Giả thuyết H_0: $\mu = 50$, đối thuyết H_1: $\mu < 50$.

Mức ý nghĩa $\alpha = 0,01 \Rightarrow t_{0,01}(24) = 2,4922$. Miền bác bỏ $W_\alpha = \{g < -2,4922\}$

$g = \dfrac{\overline{x} - \mu_0}{s}\sqrt{n} = \dfrac{49,25-50}{0,81}\sqrt{25} = -4,6296 \in W_\alpha$. Vậy nghi ngờ trên đúng.

6.4. Kiểm định giả thuyết về phương sai của đám đông

Bài toán: Giả sử đám đông X có phân phối chuẩn $N(\mu,\sigma^2)$ phương sai $V(X)=\sigma^2$ chưa biết. Với mức ý nghĩa α, kiểm định giả thuyết H_0: $\sigma^2 = \sigma_0^2$ (σ_0^2 đã biết)

- Nếu μ đã biết thì $G = \dfrac{nS^2}{\sigma_0^2} \sim \chi^2(n)$

+ Khi H_1: $\sigma^2 \neq \sigma_0^2$ miền bác bỏ $W_\alpha = \{g < \chi^2_{1-\frac{\alpha}{2}}(n)$ hoặc $g > \chi^2_{\frac{\alpha}{2}}(n)\}$

+ Khi H_1: $\sigma^2 > \sigma_0^2$ miền bác bỏ $W_\alpha = \{ g > \chi^2_\alpha(n) \}$

+ Khi H_1: $\sigma^2 < \sigma_0^2$ miền bác bỏ $W_\alpha = \{ g < \chi^2_{1-\alpha}(n) \}$

- Nếu μ chưa biết thì $G = \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

+ Khi H_1: $\sigma^2 \neq \sigma_0^2$ miền bác bỏ $W_\alpha = \{g < \chi^2_{1-\frac{\alpha}{2}}(n\text{-}1)$ hoặc $g > \chi^2_{\frac{\alpha}{2}}(n\text{-}1)\}$

+ Khi H_1: $\sigma^2 > \sigma_0^2$ miền bác bỏ $W_\alpha = \{ g > \chi^2_\alpha(n\text{-}1) \}$

+ Khi H_1: $\sigma^2 < \sigma_0^2$ miền bác bỏ $W_\alpha = \{ g < \chi^2_{1-\alpha}(n\text{-}1) \}$

Ví dụ 3: Trọng lượng của một loại sản phẩm do một máy sản xuất ra là một biến ngẫu nhiên có phân phối chuẩn. Nghi ngờ về độ đồng đều về trọng lượng sản phẩm có xu hướng giảm sút người ta cân thử 12 sản phẩm được (đơn vị: gam):

450	448	449	452
450	449	451	450
448	451	449	450

Với mức ý nghĩa 5% hãy kết luận về nghi ngờ trên biết rằng bình thường phương sai của trọng lượng sản phẩm là 1,2 $(\text{gam})^2$.

<u>Giải</u>

X	448	449	450	151	452
Tần số	2	3	4	2	1

Ta có $\overline{x} = 449{,}75$; $\overline{x^2} = 202276{,}4167$; $s^2 = 1{,}4773$

Giả thuyết H_0: $\sigma^2 = 1{,}2$, đối thuyết H_1: $\sigma^2 > 1{,}2$.

Mức ý nghĩa $\alpha = 0{,}05 \Rightarrow \chi^2_{0,05}(11) = 19{,}6751$.

Miền bác bỏ $W_\alpha = \{ g > 19{,}6751 \}$

$g = \dfrac{(n-1)s^2}{\sigma_0^2} = \dfrac{11.1{,}4773}{1{,}2} = 13{,}5419 \notin W_\alpha$. Chấp nhận H_0

BÀI TẬP CHƯƠNG 6

<u>Câu 1:</u> Lô hàng là đủ tiêu chuẩn xuất khẩu nếu tỉ lệ phế phẩm của nó không vượt quá 2%. Kiểm tra ngẫu nhiên 500 sản phẩm của lô hàng thấy có 15 phế phẩm. Với mức ý nghĩa 5% lô hàng có được xuất khẩu không?

<u>Câu 2:</u> Trọng lượng của sản phẩm A (đơn vị: kg) là biến ngẫu nhiên có phân phối chuẩn, người ta kiểm tra ngẫu nhiên trọng lượng của 21 sản phẩm này được kết quả:

35	35	34	36	34	34	35
33	36	35	35	33	34	36
36	36	35	36	37	35	35

a) Lập bảng phân phối tần số thực nghiệm, tính trung bình mẫu thực nghiệm, tính phương sai mẫu hiệu chỉnh thực nghiệm.

b) Trọng lượng trung bình của loại thiết bị A theo quy định là 36 kg, nghi ngờ trọng lượng trung bình của loại thiết bị A giảm so với quy định. Với mức ý nghĩa 1% hãy cho kết luận về nghi ngờ trên.

c) Nghi ngờ về độ đồng đều về trọng lượng sản phẩm có xu hướng giảm sút so với quy định. Với mức ý nghĩa 5% hãy cho kết luận về nghi ngờ trên biết phương sai của trọng lượng sản phẩm A theo quy định là 1 $(kg)^2$.

<u>Câu 3:</u> Trọng lượng của các bao gạo là biến ngẫu nhiên có phân phối chuẩn với trọng lượng trung bình bình thường là 50kg. Nghi ngờ máy đóng bao gạo làm việc không bình thường làm cho trọng lượng trung bình của các bao gạo có xu hướng giảm sút. Người ta cân thử ngẫu nhiên 40 bao được khối lượng như sau:

49	50	49	48	50	51	48	49	50	50
50	49	49	50	49	48	50	51	49	49
50	49	50	49	48	50	50	50	51	50
51	48	49	49	50	51	48	49	50	50

a) Lập bảng phân phối tần số thực nghiệm, tính trung bình mẫu thực nghiệm, tính phương sai mẫu hiệu chỉnh thực nghiệm.

b) Với mức ý nghĩa 1% hãy cho kết luận về nghi ngờ trên.

<u>Câu 4:</u> Mức hao phí xăng của một loại ô tô chạy từ A đến B là một biến ngẫu nhiên có phân bố chuẩn, có trung bình là 50 lít. Đoạn đường được xử lý lại, người ta cho rằng mức hao phí xăng trung bình giảm xuống. Quan sát ngẫu nhiên 50 ô tô cùng loại, người ta thu được số liệu sau

Mức hao phí X	48,5 – 49	49 – 49,5	49,5 – 50	50 – 50,5	50,5 – 51
Số chuyến	5	15	15	10	5

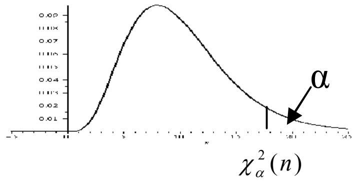
Hãy kết luận về ý kiến trên với mức ý nghĩa 5%.

Phụ lục 1: Hàm Laplace $\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

X	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986

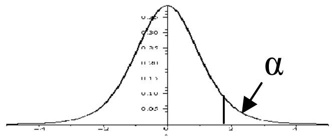
x	**3.0**	**3.1**	**3.2**	**3.3**	**3.4**
φ(x)	0.49865010	0.49903240	0.49931286	0.49951658	0.49966307
x	**3.5**	**3.6**	**3.7**	**3.8**	**3.9**
φ(x)	0.49976737	0.49984089	0.49989220	0.49992765	0.49995190
x	**4.0**	**4.5**	**5.0**		
φ(x)	0.49996833	0.49999660	0.49999971		

Phụ lục 2: $X \sim \chi^2(n)$; $P(X > \chi^2_\alpha(n))=\alpha$



n\α	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.00004	0.0002	0.0010	0.0039	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	5.9915	7.3778	9.2103	10.5966
3	0.0717	0.1148	0.2158	0.3518	7.8147	9.3484	11.3449	12.8382
4	0.2070	0.2971	0.4844	0.7107	9.4877	11.1433	13.2767	14.8603
5	0.4117	0.5543	0.8312	1.1455	11.0705	12.8325	15.0863	16.7496
6	0.6757	0.8721	1.2373	1.6354	12.5916	14.4494	16.8119	18.5476
7	0.9893	1.2390	1.6899	2.1673	14.0671	16.0128	18.4753	20.2777
8	1.3444	1.6465	2.1797	2.7326	15.5073	17.5345	20.0902	21.9550
9	1.7349	2.0879	2.7004	3.3251	16.9190	19.0228	21.6660	23.5894
10	2.1559	2.5582	3.2470	3.9403	18.3070	20.4832	23.2093	25.1882
11	2.6032	3.0535	3.8157	4.5748	19.6751	21.9200	24.7250	26.7568
12	3.0738	3.5706	4.4038	5.2260	21.0261	23.3367	26.2170	28.2995
13	3.5650	4.1069	5.0088	5.8919	22.3620	24.7356	27.6882	29.8195
14	4.0747	4.6604	5.6287	6.5706	23.6848	26.1189	29.1412	31.3193
15	4.6009	5.2293	6.2621	7.2609	24.9958	27.4884	30.5779	32.8013
16	5.1422	5.8122	6.9077	7.9616	26.2962	28.8454	31.9999	34.2672
17	5.6972	6.4078	7.5642	8.6718	27.5871	30.1910	33.4087	35.7185
18	6.2648	7.0149	8.2307	9.3905	28.8693	31.5264	34.8053	37.1565
19	6.8440	7.6327	8.9065	10.1170	30.1435	32.8523	36.1909	38.5823
20	7.4338	8.2604	9.5908	10.8508	31.4104	34.1696	37.5662	39.9968
21	8.0337	8.8972	10.2829	11.5913	32.6706	35.4789	38.9322	41.4011
22	8.6427	9.5425	10.9823	12.3380	33.9244	36.7807	40.2894	42.7957
23	9.2604	10.1957	11.6886	13.0905	35.1725	38.0756	41.6384	44.1813
24	9.8862	10.8564	12.4012	13.8484	36.4150	39.3641	42.9798	45.5585
25	10.5197	11.5240	13.1197	14.6114	37.6525	40.6465	44.3141	46.9279
26	11.1602	12.1981	13.8439	15.3792	38.8851	41.9232	45.6417	48.2899
27	11.8076	12.8785	14.5734	16.1514	40.1133	43.1945	46.9629	49.6449
28	12.4613	13.5647	15.3079	16.9279	41.3371	44.4608	48.2782	50.9934
29	13.1211	14.2565	16.0471	17.7084	42.5570	45.7223	49.5879	52.3356
30	13.7867	14.9535	16.7908	18.4927	43.7730	46.9792	50.8922	53.6720

Phụ lục 3: $X \sim t(n)$; $P(X > t_\alpha(n)) = \alpha$



n\α	0.005	0.01	0.02	0.025	0.05	0.1
1	63.6567	31.8205	15.8945	12.7062	6.3138	3.0777
2	9.9248	6.9646	4.8487	4.3027	2.9200	1.8856
3	5.8409	4.5407	3.4819	3.1824	2.3534	1.6377
4	4.6041	3.7469	2.9985	2.7764	2.1318	1.5332
5	4.0321	3.3649	2.7565	2.5706	2.0150	1.4759
6	3.7074	3.1427	2.6122	2.4469	1.9432	1.4398
7	3.4995	2.9980	2.5168	2.3646	1.8946	1.4149
8	3.3554	2.8965	2.4490	2.3060	1.8595	1.3968
9	3.2498	2.8214	2.3984	2.2622	1.8331	1.3830
10	3.1693	2.7638	2.3593	2.2281	1.8125	1.3722
11	3.1058	2.7181	2.3281	2.2010	1.7959	1.3634
12	3.0545	2.6810	2.3027	2.1788	1.7823	1.3562
13	3.0123	2.6503	2.2816	2.1604	1.7709	1.3502
14	2.9768	2.6245	2.2638	2.1448	1.7613	1.3450
15	2.9467	2.6025	2.2485	2.1314	1.7531	1.3406
16	2.9208	2.5835	2.2354	2.1199	1.7459	1.3368
17	2.8982	2.5669	2.2238	2.1098	1.7396	1.3334
18	2.8784	2.5524	2.2137	2.1009	1.7341	1.3304
19	2.8609	2.5395	2.2047	2.0930	1.7291	1.3277
20	2.8453	2.5280	2.1967	2.0860	1.7247	1.3253
21	2.8314	2.5176	2.1894	2.0796	1.7207	1.3232
22	2.8188	2.5083	2.1829	2.0739	1.7171	1.3212
23	2.8073	2.4999	2.1770	2.0687	1.7139	1.3195
24	2.7969	2.4922	2.1715	2.0639	1.7109	1.3178
25	2.7874	2.4851	2.1666	2.0595	1.7081	1.3163
26	2.7787	2.4786	2.1620	2.0555	1.7056	1.3150
27	2.7707	2.4727	2.1578	2.0518	1.7033	1.3137
28	2.7633	2.4671	2.1539	2.0484	1.7011	1.3125
29	2.7564	2.4620	2.1503	2.0452	1.6991	1.3114
30	2.7500	2.4573	2.1470	2.0423	1.6973	1.3104

TÀI LIỆU THAM KHẢO

1. Lê Sĩ Đồng, *Xác suất thống kê và ứng dụng*, NXB Giáo dục, 2004.

2. Đặng Hùng Thắng, *Mở đầu về Lý thuyết xác suất và các ứng dụng*, NXB Giáo dục, 2008.

3. Đặng Hùng Thắng, *Thống kê và ứng dụng*, NXB Giáo dục, 2008.

4. Đinh Văn Gắng, *Bài tập xác suất và thống kê*, NXB Giáo dục, 2003.

5. Nguyễn Văn Cao, Trần Thái Ninh, *Giáo trình lí thuyết xác suất và thống kê toán*, NXB Thống kê, Hà Nội, 2005.

MỤC LỤC

Trang

Chương 1. Các khái niệm cơ bản trong lý thuyết xác suất 1

Chương 2. Biến ngẫu nhiên..8

Chương 3. Các phân phối xác suất thường dùng 14

Chương 4. Lý thuyết mẫu..19

Chương 5. Lý thuyết ước lượng...23

Chương 6. Kiểm định giả thuyết thống kê 30

Phụ lục 1 33

Phụ lục 2 34

Phụ lục 3 35

Tài liệu tham khảo 36